The Dissertation Committee for Hannah Pearl Flath
certifies that this is the approved version of the following dissertation:

Hessian-based response surface approximations for uncertainty quantification in large-scale statistical inverse problems, with applications to groundwater flow

Committee:

Omar Ghattas, Supervisor

Todd Arbogast

Clint Dawson

Marc Hesse

Serge Prudhomme

Hessian-based response surface approximations for uncertainty quantification in large-scale statistical inverse problems, with applications to groundwater flow

by

Hannah Pearl Flath, B.A., M.S.C.A.M.

DISSERTATION

Presented to the Faculty of the Graduate School of
The University of Texas at Austin
in Partial Fulfillment
of the Requirements
for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN
August 2013

Dedicated to my father, Daniel E. Flath, who taught me to love math.

Acknowledgments

The list of acknowledgments and thank-yous is always long when one pours heart and soul into research and development over any time – this list must start with my advisor, Omar Ghattas, without whom a successful finish would not have been possible, and my undergraduate advisor, David Keyes, whose inspiration started this journey. I also wish to thank the CCGO research group for many, many helpful discussions and extensive advice. Many thanks to my dissertation committee for all their advice and guidance, and of course there are my fellows and dear friends at ICES who spent countless hours in discussion offering friendship, support, and advice. I thank the members of the UT Aikido Club for their companionship, their enthusiasm, and helping me get up and move. I also thank my many wonderful friends in the Austin area who reminded me that there is life after graduate school. No list of thanks would be complete without Elizabeth Hicks, whose emails and support were invaluable. Finally, thank you so very much to my family – Mom, Dad, brothers, grandparents, aunts, uncles, and cousins who believed in me through it all.

Hessian-based response surface approximations for uncertainty quantification in large-scale statistical inverse problems, with applications to groundwater flow

Hannah Pearl Flath, Ph.D.
The University of Texas at Austin, 2013

Supervisor: Omar Ghattas

Subsurface flow phenomena characterize many important societal issues in energy and the environment. A key feature of these problems is that subsurface properties are uncertain, due to the sparsity of direct observations of the subsurface. The Bayesian formulation of this inverse problem provides a systematic framework for inferring uncertainty in the properties given uncertainties in the data, the forward model, and prior knowledge of the properties. We address the problem: given noisy measurements of the head, the pdf describing the noise, prior information in the form of a pdf of the hydraulic conductivity, and a groundwater flow model relating the head to the hydraulic conductivity, find the posterior probability density function (pdf) of the parameters describing the hydraulic conductivity field. Unfortunately, conventional sampling of this pdf to compute statistical moments is intractable for problems governed by large-scale forward models and high-dimensional parameter spaces.

We construct a Gaussian process surrogate of the posterior pdf based on Bayesian interpolation between a set of "training" points. We employ a greedy algorithm to find the training points by solving a sequence of optimization problems where each new training point is placed at the maximizer of the error in the approximation. Scalable Newton optimization methods solve this "optimal" training point problem. We tailor the Gaussian process surrogate to the curvature of the underlying posterior pdf according to the Hessian of the log posterior at a subset of training points, made computationally tractable by a low-rank approximation of the data misfit Hessian. A Gaussian mixture approximation of the posterior is extracted from the Gaussian process surrogate, and used as a proposal in a Markov chain Monte Carlo method for sampling both the surrogate as well as the true posterior. The Gaussian process surrogate is used as a first stage approximation in a two-stage delayed acceptance MCMC method.

We provide evidence for the viability of the low-rank approximation of the Hessian through numerical experiments on a large scale atmospheric contaminant transport problem and analysis of an infinite dimensional model problem. We provide similar results for our groundwater problem. We then present results from the proposed MCMC algorithms.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Subsurface flow and transport phenomena characterize many important societal issues in energy and the environment, such as groundwater contamination, management of nuclear waste repositories, water resources management, carbon sequestration, and enhanced oil recovery. A key feature of these problems is that subsurface properties (such as permeability and porosity) are uncertain, due to the sparsity of direct observations of the subsurface. Subsurface properties can be inferred from observations via solution of an inverse problem governed by the forward equations of subsurface flow and transport. There is a long history of the application of classical deterministic inverse methods to solve such PDE-constrained inverse problems [48]. However, the classical solution gives no indication of the confidence we have in the solution of the inverse problem, i.e., it is incapable of quantifying uncertainties in the subsurface properties inferred from noisy data and an imperfect model.

The Bayesian formulation of the statistical inverse problem, on the other hand, provides a systematic framework for inferring uncertainty in the properties given uncertainties in the data, the forward model, and prior knowledge of the properties. In the Bayesian approach (see [38, 52]), the solution of the inverse problem is a probability density function (pdf) of parameters characterizing the properties, the so-called posterior pdf. Unfortunately, sampling this pdf (for example using Markov chain Monte Carlo (MCMC) methods) to compute statistical moments such as mean, variance, etc., is essentially intractable for statistical inverse problems governed by large-scale forward models (such as the PDEs of subsurface flow) and high-dimensional parameter spaces (such as discretizations of permeability fields). Approaches to solving nonlinear statistical inverse problems must deal with this issue. The goals of this thesis are to develop methods that address the difficulty of sampling high-dimensional expensive posterior pdf's by exploiting problem structure, and to apply these methods to inference of hydraulic conductivity in groundwater flow problems.

One approach to overcoming the high cost of sampling the posterior pdf is to create a low cost surrogate parameter-to-observable map. The parameter-to-observable map is replaced with a Gaussian process response surface in [39]. Stochastic collocation methods are used to construct a generalized polynomial chaos approximation of the forward model over the support of the prior in [44]. Polynomial chaos expansions are also used to construct a surrogate model in [4]. In the proper orthogonal decomposition (POD) method, the forward problem is projected onto a reduced subspace [56]. For a comparison of stochastic collocation methods versus POD for the construction of surrogate forward models, see [37].

A second approach reduces the dimension of the parameter space, and thus sampling of the posterior pdf is no longer in a high dimensional space. In [22], the model parameters are represented with a Karhunen-Loève (K-L) expansion with the truncation point related to the decay in the spectrum of the covariance matrix. Truncated K-L expansions based on the prior random process have also been used to approximate model parameters (e.g. [45]). Another method of reduction of the dimension of the parameter space is projection of the high dimensional parameter space into a reduced subspace (e.g. [40]). In [41], the first and second approaches are combined: a reduced model for both the state space and parameter space is developed using a greedy algorithm.

A third approach guides the sampling process in order to improve the efficiency of MCMC. Two stage MCMC schemes increase the acceptance rate by using a coarse scale model, where the coarse scale model may be based on single-phase upscaling [24] or multiscale finite volume methods [25]. References [23, 22] guide MCMC with the Langevin method using gradients from coarse scale models. A local linear approximation to the forward model is used in [19] to improve the acceptance probability. Reference [35] applies an MCMC scheme wherein information from a coarse-scale model influences a fine-scale chain through swap proposals. Another two stage approach is proposed in [21], and combines a reduced order model with an adaptive delayed-acceptance algorithm to build a stochastic model of the error in the model to improve the first-stage performance. Reference [43] introduces the Stochastic Newton MCMC algorithm, which incorporates Hessian information and has been successful for large scale problems. However, it is based on a local Gaussian approximation, which results in low acceptance rates when this is a poor approximation. Our approach addresses this problem by replacing the Gaussian proposal of the stochastic Newton method with a mixture of Gaussians, thereby allowing the proposal to better capture multi-modality of the posterior. Like the stochastic Newton method, these Gaussians are tailored to local inverse Hessian information.

While these methods have been successful in permitting solution of Bayesian inverse problems in modest dimensions, they do not allow scalability to high dimensions. This work will contribute to the field by approximating the posterior pdf directly, and addressing issues of scaling through tailoring state-of-the-art optimization algorithms. The prohibitive cost of applying standard sampling methods to such large-scale posterior pdf's stems from their view of the map from parameters to observations (which involves solution of the forward problem) as a black box. That is, current methods for statistical inverse problems do not exploit advances in large-scale deterministic inverse problems such as fast computation of derivatives of this map and Hessian-vector products via adjoint methods. The exception is [23, 22], which uses coarse-scale gradients in the Langevin method to guide sampling. In this thesis, we develop methods for approximating posterior pdf's that employ Gaussian process models that are informed by Hessian information, computed efficiently through adjoint methods and low-rank spectral approximations. We will then apply these methods to the inverse problem of inferring hydraulic conductivity of the subsurface from measurements of piezometric head and a single-phase groundwater flow forward model.

We base the posterior pdf approximation on Bayesian interpolation between a set of “training” points of the posterior pdf, which provides not only an approximation but also an estimate of the uncertainty in the approximation. Ensuring a good approximation to the posterior pdf based on a randomly chosen set of training points requires a number of evaluations which increases with the dimension of the space, and becomes intractable in high-dimensional parameter spaces. Instead we employ a greedy algorithm [17], which transforms the choice of training points into a series of optimization problems where each new training point is placed at the maximizer of the square of the error in the approximation. We take advantage of scalable optimization methods such as Inexact Newton-Conjugate Gradient, in addition to adjoint-based gradients of both the posterior pdf and the response surface approximation, to solve this “optimal” training point problem.

For the construction of the response surface approximation through Bayesian interpolation, we choose a Gaussian process prior. This requires a choice of a prior mean and covariance (we assume no noise, since the value of the posterior pdf is evaluated exactly). We choose the prior mean equal to zero, since a high-dimensional pdf will be near zero in most regions. To improve the accuracy of the approximation we tailor the response surface to the curvature of the underlying posterior pdf (through the choice of the prior covariance) according to the Hessian of the log of the posterior pdf at each of the training points. This Hessian is related to the Hessian of a discretization of the objective function for the classical deterministic formulation as a constrained optimization problem, and thus the response surface approximation incorporates information from the underlying forward model into the approximation. We build a second response surface from only those points with Hessian information: combining them into a Gaussian mixture. While this is a less accurate approximation (as it is not an interpolation), it is easy to draw samples from this distribution. Specifics of these methods are given in Chapter 2.

We then draw upon these Hessian-based response surfaces to propose three MCMC algorithms to calculate desired properties of the posterior pdf such as the mean and variance. These three algorithms share the upfront cost of building the response surfaces, but differ in the additional cost to obtain a sample. One option, the most accurate and most costly, samples the posterior pdf with the Gaussian mixture as a proposal in the Metropolis-Hastings method. For a cheaper (but less accurate) option, we sample the Gaussian process response surface as a surrogate for the posterior pdf, with the Gaussian mixture as a proposal. The third option is a two-stage delayed acceptance algorithm: a first-stage accepts or rejects a point based on the Gaussian process response surface, and accepted points face a second-stage test based on the exact posterior pdf. This combines accuracy from sampling the true posterior pdf while reducing the cost relative to the single-stage method. All three rely on access to information about the Hessian.

However, construction of the Hessian formally requires n forward and adjoint solves, where n is the number of parameters. Scalability to high-dimensional parameter spaces then requires a low-rank approximation of the Hessian. We note that the Hessian is composed of two terms: the Hessian of the data misfit and the Hessian of the prior term. For many ill-posed inverse problems, the Hessian of the data misfit behaves like the discretization of a compact operator. The range space thus is effectively finite-dimensional, and the eigenvalues decay, often rapidly, to zero. We can exploit this structure to construct fast algorithms for approximating the Hessian [26]. Detailed exposition of the low-rank approximation is in Chapter 4 with analytic and numerical evidence for a nonlinear inverse problem, while derivation of the Hessian for this application is in Chapter 3.

Thus the goal, stated at a general level, is to combine the Bayesian framework for inverse problems [52], Gaussian process response surface approximations [34, 39], adaptive sampling [17, 15], and low-rank Hessian approximations [26] to provide a solution to the statistical inverse problem governed by groundwater flow. This is an intrusive approach which takes advantage of information from the Hessian about the structure of the underlying problem to construct an approximation to the posterior pdf. The method addresses issues of scale through exploiting the structure of the Hessian and tailored state-of-the-art optimization algorithms. Chapter 2 details the Bayesian framework and the construction of the response surface approximation, while Chapter 3 introduces the application to groundwater flow and provides a two parameter example to illustrate the algorithms presented in Chapter 2. Chapter

4 describes the low-rank approximation in detail, Chapter 5 demonstrates a linear convection-diffusion inverse problem, and Chapter 6 demonstrates the algorithm for a nonlinear groundwater flow problem.

Chapter 2

Uncertainty quantification for nonlinear inverse problems

Here we present the Bayesian framework for statistical inverse problems beginning with the general case of Bayes' theorem and continuing with the special case of Gaussian noise and prior uncertainties. Due to the nonlinearity of the inverse problem, the resulting posterior pdf will not necessarily be Gaussian. Our focus here is on inverse problems that are governed by large-scale forward models, as result from appropriate discretizations of PDEs. Unfortunately, for large-scale statistical inverse problems (as exemplified by the groundwater flow problem we target here), a conventional approach of directly sampling this posterior pdf is entirely prohibitive.

Hence, in §2.2, we present an algorithm to provide an approximation of the posterior pdf through a series of optimization problems. Bayesian interpolation based on these observations then provides both an approximation of the posterior pdf and an estimate of its accuracy. This method is an effective way of selecting training points to provide an accurate picture of the posterior pdf in the context of high-dimensional parameter spaces. To construct an interpolation that matches the true posterior pdf well, we shape the approximation according to the Hessian of the negative log of the posterior pdf at a subset of the training points. To provide scalability to high-dimensional parameter spaces, in §4 we present a fast method for estimating the Hessian based on a low rank spectral estimate that approximates its structure. §3.2 provides a derivation of the Hessian for the groundwater flow problem.

In this chapter, we approach the problem of a nonlinear Bayesian statistical inverse problem in stages. First we approach the statistical inverse problem through the Bayesian framework to obtain an expression for the posterior pdf. Second, we discuss how to choose training points at which to evaluate the posterior pdf according to a series of optimization problems. We then construct an approximation and an estimate of its accuracy with a Bayesian interpolation method by taking advantage of information provided by the Hessian through the interpolation prior. This information also allows construction of a Gaussian mixture proposal density.

2.1 General Bayesian Formulation

Non-uniqueness is a central feature of ill-posed inverse problems: multiple values of the parameters may be consistent with the observations. The least-squares minimization approach to ill-posed inverse problems invokes a so-called regularization term to effectively select among the multiple parameter values the one that has largest regularity (in an appropriate norm), resulting in a single—deterministic—estimate of the unknown parameters (see, e.g., [55]). A Bayesian estimate of the unknown, on the other hand, is a probability density that suggests the credibility of any given point estimate (see, e.g., [38, 52]). In the Bayesian approach, we view all parameters as random variables and write the parameter-to-observable map $g : \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}^m$ as

$$Y = g(X, E),$$

where X, Y, and E are random variables. The variable $\boldsymbol{x} \in \mathbb{R}^n$ is a realization of the random variable X representing the vector of model parameters to be recovered, $\boldsymbol{e} \in \mathbb{R}^k$ is a realization of the random variable E representing the vector of errors (due to both model errors and observation noise), and $\boldsymbol{y} \in \mathbb{R}^m$ is the realization of random variable Y representing the vector of observables, with $\boldsymbol{y}_{\text{obs}}$ the actual observation values. We choose the following probability density functions (pdf's): the probability density $\pi_{\text{noise}} : \mathbb{R}^k \to \mathbb{R}$, which describes the modeling error and observation noise; the prior probability density $\pi_{\text{prior}} : \mathbb{R}^n \to \mathbb{R}$, which describes additional information about the parameters X; and the likelihood function $\pi(\boldsymbol{y}|\boldsymbol{x})$, which describes the relationship between the observables $\boldsymbol{y}$ and the unknown model parameters $\boldsymbol{x}$.

From Bayes' theorem, the prior probability density, the likelihood function, and the data can be combined to form the posterior probability density $\pi_{\text{post}} : \mathbb{R}^n \to \mathbb{R}$ of the model parameters X:

$$\begin{aligned}\pi_{\text{post}}(\boldsymbol{x}) &:= \pi(\boldsymbol{x}|\boldsymbol{y}_{\text{obs}}) \\ &= \frac{\pi_{\text{prior}}(\boldsymbol{x})\,\pi(\boldsymbol{y}_{\text{obs}}|\boldsymbol{x})}{\pi(\boldsymbol{y}_{\text{obs}})} \\ &\propto \pi_{\text{prior}}(\boldsymbol{x})\,\pi(\boldsymbol{y}_{\text{obs}}|\boldsymbol{x}).\end{aligned}$$

To generate the likelihood function, we use the pdf $\pi_{\text{noise}}(\boldsymbol{e})$. Here we will assume additive noise, so that the parameter-to-observable map is

$$Y = f(X) + E,$$

where $f : \mathbb{R}^n \to \mathbb{R}^m$ and the noise $E \in \mathbb{R}^m$ reflects both the modeling error of f and observation noise. Thus $E = Y - f(X)$. We assume that X and E are statistically independent (see e.g., [38, p.56], for generalizations). Therefore,

$$\pi_{\text{noise}}(\boldsymbol{e}) = \pi_{\text{noise}}(\boldsymbol{y}_{\text{obs}} - f(\boldsymbol{x})),$$

so that Bayes' theorem states

$$\pi_{\text{post}}(\boldsymbol{x}) \propto \pi_{\text{prior}}(\boldsymbol{x})\, \pi_{\text{noise}}(\boldsymbol{y}_{\text{obs}} - f(\boldsymbol{x})).$$

In the case that the prior probability density of X and the probability density of error E are both Gaussian, further simplifications can be made. The algorithms presented in this dissertation, however, are not dependent on these simplifications. The prior and noise pdf's can be written in the form

$$\begin{aligned} \pi_{\text{prior}}(\boldsymbol{x}) &\propto \exp\left(-\frac{1}{2}(\boldsymbol{x} - \bar{\boldsymbol{x}}_{\text{prior}})^T \boldsymbol{\Gamma}_{\text{prior}}^{-1}(\boldsymbol{x} - \bar{\boldsymbol{x}}_{\text{prior}})\right), \\ \pi_{\text{noise}}(\boldsymbol{e}) &\propto \exp\left(-\frac{1}{2}(\boldsymbol{e} - \bar{\boldsymbol{e}})^T \boldsymbol{\Gamma}_{\text{noise}}^{-1}(\boldsymbol{e} - \bar{\boldsymbol{e}})\right), \end{aligned}$$

where $\bar{\boldsymbol{x}}_{\text{prior}} \in \mathbb{R}^n$ is the mean of the prior pdf, $\bar{\boldsymbol{e}} \in \mathbb{R}^m$ is the mean of the noise pdf, $\boldsymbol{\Gamma}_{\text{prior}} \in \mathbb{R}^{n\times n}$ is the covariance matrix of the prior pdf, and $\boldsymbol{\Gamma}_{\text{noise}} \in \mathbb{R}^{m\times m}$ is the covariance matrix of the noise pdf. Restating Bayes' theorem with these Gaussian uncertainties, we find that

$$\pi_{\text{post}}(\boldsymbol{x}) \propto \exp\left(-\frac{1}{2}\|\boldsymbol{x} - \bar{\boldsymbol{x}}_{\text{prior}}\|^2_{\boldsymbol{\Gamma}_{\text{prior}}^{-1}} - \frac{1}{2}\|\boldsymbol{y}_{\text{obs}} - f(\boldsymbol{x}) - \bar{\boldsymbol{e}}\|^2_{\boldsymbol{\Gamma}_{\text{noise}}^{-1}}\right). \tag{2.1}$$

As is clear from this expression, despite the choice of prior and noise probability densities as Gaussian, the posterior pdf need not be Gaussian due to the nonlinearity of $f(\boldsymbol{x})$. Note that the log of this expression is related to the discretization of the least-squares optimization problem, and therefore the maximum *a posteriori* (MAP) point of the pdf is the solution of the classical deterministic inverse problem.

2.2 Approximation of $\pi_{\mathbf{post}}(\boldsymbol{x})$

While we now have an expression for the posterior pdf, sampling the posterior conventionally is impractical due to the high dimensional nature of the parameter space and the computational cost of solving the PDE model. To address this problem, we construct a response surface approximation $\hat{\pi}_{\text{post}}(\boldsymbol{x})$ to the posterior pdf $\pi_{\text{post}}(\boldsymbol{x})$. This response surface approximation is essentially an interpolation between training points at which $\pi_{\text{post}}(\boldsymbol{x})$ is evaluated, developed through the Bayesian framework with an accompanying estimate of the error in the approximation. Three key points arise: first we advocate transforming the choice of training points for a high-dimensional parameter space into a series of optimization problems. Second, the response surface approximation should be shaped by the Hessian $\nabla^2_{\boldsymbol{x}}(-\log \pi_{\text{post}}(\boldsymbol{x}))$ of the underlying statistical inverse problem so that it may respect the local structure of the underlying pdf. This Hessian information may be reused to construct a Gaussian mixture proposal density and to indicate good search directions for the greedy algorithm. Third, we propose the construction of a low-rank approximation of the Hessian to allow scalability of the response surface approximation in high-dimensional parameter spaces; this topic is deferred to Chapter 4. A step-by-step illustration of the algorithm presented in this section for a two parameter groundwater flow problem is given in §3.3.

2.2.1 Adaptive sampling algorithm

The adaptive algorithm described in this section is general and independent of the specific choice of interpolation; the choice of interpolation will be described later in §2.2.2. As discussed earlier, we wish to construct a response surface approximation based on interpolating between training points: *value points* at which the true value of the posterior pdf is known and *shape points* at which we have additional information about the Hessian $\nabla^2_{\boldsymbol{x}}(-\log \pi_{\text{post}}(\boldsymbol{x}))$. However, ensuring the accuracy of this approximation of $\pi_{\text{post}}(\boldsymbol{x})$ for a randomly chosen set of training points requires a prohibitive number of points for high-dimensional parameter spaces. Let the set of training points after j greedy iterations be $\boldsymbol{M}_j = \{\boldsymbol{x}_i\}_{i=1}^{n_j} = \boldsymbol{M}_j^{\text{shape}} \cup \boldsymbol{M}_j^{\text{value}}$. Let the approximation (based on these training points) be $\hat{\pi}_{\text{post}}(\boldsymbol{x}, \boldsymbol{M}_j)$. The best possible set of training points is a result of the optimization problem to minimize the global error

$$\min_{\boldsymbol{M}_j} \|\pi_{\text{post}}(\boldsymbol{x}) - \hat{\pi}_{\text{post}}(\boldsymbol{x}, \boldsymbol{M}_j)\|.$$

Evaluation of this objective requires an integral over n-dimensional space, which is intractable for a high dimensional parameter space. Instead, we use the greedy algorithm as a heuristic for improving our set of training points.

With a greedy algorithm, the "optimal" training point problem is converted to a series of optimization problems. Each new shape point is a result of the problem

$$\max_{\boldsymbol{\gamma}} \; (\pi_{\text{post}}(\boldsymbol{\gamma}) - \hat{\pi}_{\text{post}}(\boldsymbol{\gamma}))^2,$$

so that the new shape point is placed where the error in the approximation is largest. This takes advantage of scalable large-scale optimization methods such as Inexact Newton-Conjugate Gradients. Additional value points are added along the path of the optimization algorithm, incorporating readily available information to improve the accuracy of the approximation.

We now describe the adaptive sampling algorithm in more detail. Let the set of training points for the kth iteration be $\boldsymbol{M}_k$. Let $\mathcal{G}(\boldsymbol{x}, \boldsymbol{M}_k)$ be the mean-squared error function at the kth iteration:

$$\mathcal{G}(\boldsymbol{x}, \boldsymbol{M}_k) = (\pi_{\text{post}}(\boldsymbol{x}) - \hat{\pi}_{\text{post}}(\boldsymbol{x}, \boldsymbol{M}_k))^2.$$

Then, at each iteration, the steps of the adaptive sampling algorithm are:

1. Choose an initial guess for the optimization problem in this greedy step.
2. Solve the optimization problem

   $$\boldsymbol{x}_{\text{new}} = \arg\max_{\boldsymbol{x}} \mathcal{G}(\boldsymbol{x}, \boldsymbol{M}_k)$$

   to find the point $\boldsymbol{x}_{\text{new}}$ where the error $\mathcal{G}(\boldsymbol{x}, \boldsymbol{M}_k)$ is maximized.
3. Update the set of value points $\boldsymbol{M}_{k+1}^{value}$ to include points along the path of the optimization where $\pi_{\text{post}}(\boldsymbol{x})$ was evaluated during the process and is thus available without extra cost.

4. If $\mathcal{G}(\boldsymbol{x}_{\text{new}}, \boldsymbol{M}_k) < \varepsilon$, where ε is the tolerance, then do not add a shape point. If $\pi_{\text{post}}(\boldsymbol{x}_{\text{new}})$ is approximately zero, then a value point for interpolation is enough, and we do not add a shape point. Otherwise, update the set of shape points of the posterior pdf so that $\boldsymbol{M}_{k+1}^{shape} = \{\boldsymbol{x}_i\}_{i=1}^{s_{k+1}}$, to include $\boldsymbol{x}_{\text{new}}$.
5. Update the approximation $\hat{\pi}_{\text{post}}(\boldsymbol{x}, \boldsymbol{M}_{k+1})$
6. If the approximation is sufficiently accurate, then terminate. The desired approximation is then the response surface approximation $\hat{\pi}_{\text{post}}(\boldsymbol{x})$ based on the n_{k+1} training points already chosen. Otherwise, continue.
7. Return to step 1.

"Sufficiently accurate" for terminating the greedy algorithm is difficult to determine, given that we cannot calculate a global error measure. However, as a heuristic, we consider the approximation complete after several greedy iterations in a row return an error less than the tolerance. The number of greedy iterations may also be constrained by cost.

2.2.2 Bayesian interpolation

We now discuss the construction of the approximation $\hat{\pi}_{\text{post}}(\boldsymbol{x}, \boldsymbol{M}_k)$ based on the values of the posterior pdf $\{\pi_{\text{post}}(\boldsymbol{x}_i)\}_{i=1}^{n_k}$ at the training points $\boldsymbol{M}_k = \{\boldsymbol{x}_i\}_{i=1}^{n_k}$. In addition, we want an estimate of the uncertainty in the approximation. Again, we turn to the Bayesian framework, this time for a statistical interpolation approach. However, instead of probability densities over a finite number of parameters as discussed in previous sections, we consider densities over functions by using Gaussian processes. By definition, the probability density of a random function $\pi_{\text{post}}(\boldsymbol{x})$ is a Gaussian process if for any set of points $\{\boldsymbol{x}_i\}_{i=1}^{n}$ for any n, the marginal density $\pi(\pi_{\text{post}}(\boldsymbol{x}_1), \pi_{\text{post}}(\boldsymbol{x}_2), \ldots, \pi_{\text{post}}(\boldsymbol{x}_n))$ is a multivariate Gaussian.

According to the Bayesian framework, the problem statement is: given evaluations $\{\pi_{\text{post}}(\boldsymbol{x}_i)\}_{i=1}^{n_k}$ of the function $\pi_{\text{post}}(\boldsymbol{x})$ and a Gaussian process prior $\pi_{\text{inter}}(\boldsymbol{x}) = \mathcal{GP}(\mu_{\text{inter}}(\boldsymbol{x}), \Gamma(\boldsymbol{x}))$ describing prior knowledge about $\pi_{\text{post}}(\boldsymbol{x})$, find the posterior interpolation pdf that approximates $\pi_{\text{post}}(\boldsymbol{x})$. Since $\pi_{\text{post}}(\boldsymbol{x})$ is evaluated directly, this is a linear inverse problem. For a linear problem with Gaussian process uncertainties, Bayes' theorem results in the Gaussian process posterior

$$\pi_{\text{GPRS}}(\boldsymbol{x}) = \mathcal{GP}(\mu_{\text{GPRS}}(\boldsymbol{x}), \Gamma_{\text{GPRS}}(\boldsymbol{x}))$$

described by the expressions [50]

$$\begin{aligned}
\mu_{\text{GPRS}}(\boldsymbol{x}) &= \mu_{\text{inter}}(\boldsymbol{x}) + \mathbf{\Gamma}^T(\boldsymbol{M}_k, \boldsymbol{x})\mathbf{\Gamma}^{-1}(\boldsymbol{M}_k, \boldsymbol{M}_k)(\pi_{\text{post}}(\boldsymbol{M}_k) - \mu_{\text{inter}}(\boldsymbol{M}_k)) \\
\Gamma_{\text{GPRS}}(\boldsymbol{x}_i, \boldsymbol{x}_j) &= \Gamma(\boldsymbol{x}_i, \boldsymbol{x}_j) - \mathbf{\Gamma}^T(\boldsymbol{M}_k, \boldsymbol{x}_i)\mathbf{\Gamma}^{-1}(\boldsymbol{M}_k, \boldsymbol{M}_k)\mathbf{\Gamma}(\boldsymbol{M}_k, \boldsymbol{x}_j),
\end{aligned}$$

where $\mu_{\text{inter}}(\boldsymbol{M}_k)$ is a vector whose ith element is $\mu_{\text{inter}}(\boldsymbol{x}_i)$, $\pi_{\text{post}}(\boldsymbol{M}_k)$ is a vector whose ith element is $\pi_{\text{post}}(\boldsymbol{x}_i)$, $\mathbf{\Gamma}(\boldsymbol{M}_k, x)$ is a vector whose ith element is $\Gamma(\boldsymbol{x}_i, \boldsymbol{x})$, and $\mathbf{\Gamma}(\boldsymbol{M}_k, \boldsymbol{M}_k)$ is a matrix whose ijth element is $\Gamma(\boldsymbol{x}_i, \boldsymbol{x}_j)$.

For our interpolation problem, we choose the Gaussian process prior

$$\pi_{\text{inter}}(\boldsymbol{x}) = \mathcal{GP}(0, \Gamma_k),$$

where the (nonstationary) covariance Γ_k based on the shape points $\boldsymbol{M}_k^{\text{shape}}$ is

$$\Gamma_k(\boldsymbol{x}_i, \boldsymbol{x}_j) = \sum_{l=1}^{s_k} P(\boldsymbol{x}_l|\boldsymbol{x}_i) P(\boldsymbol{x}_l|\boldsymbol{x}_j) \exp\left(-\frac{1}{2}\|\boldsymbol{x}_i - \boldsymbol{x}_j\|_{\boldsymbol{H}_l}^2\right)$$

where

$$P(\boldsymbol{x}_l|\boldsymbol{x}_i) = \frac{\exp\left(-\frac{1}{2}\|\boldsymbol{x}_i - \boldsymbol{x}_l\|_{\boldsymbol{H}_l}^2\right)}{\sum_{p=1}^{s_k} \exp\left(-\frac{1}{2}\|\boldsymbol{x}_i - \boldsymbol{x}_p\|_{\boldsymbol{H}_p}^2\right)}$$

and

$$\boldsymbol{H}_i \approx \nabla_{\boldsymbol{x}}^2(-\log \pi_{\text{post}}(\boldsymbol{x}))\big|_{\boldsymbol{x}=\boldsymbol{x}_i}$$

is a positive definite approximation to the Hessian of the negative log of the posterior pdf (2.1). For a high-dimensional parameter space, the value of a probability density $\pi(\boldsymbol{x})$ is nearly zero over most of the domain, particularly if the probability mass is concentrated. Thus the choice of zero prior mean is appropriate. The fundamentally important step is incorporating the local structure of the underlying statistical inverse problem in the form of the Hessian into the problem of interpolating between training points. The prior covariance function Γ_k determines the shape of the interpolation between training points to match the local curvature of the posterior pdf, described by the approximation of the Hessian of the negative log of the posterior pdf at the training points $\boldsymbol{x}_i$. This is accomplished through a sum of Gaussians (shaped by the Hessian at the given training points) weighted by a membership probability to determine the relative contributions of the modes. Applications to synthetic problems in [15] suggest several orders of magnitude improvement in accuracy due to inclusion of the Hessian.

We restate the expressions for the interpolation posterior mean and variance with the chosen Gaussian process prior:

$$\begin{aligned}\mu_{\mathrm{GPRS}}(\boldsymbol{x}) &= \boldsymbol{\Gamma}_k^T(\boldsymbol{M}_k,\boldsymbol{x})\boldsymbol{\Gamma}_k^{-1}(\boldsymbol{M}_k,\boldsymbol{M}_k)\pi_{\mathrm{post}}(\boldsymbol{M}_k)\\ \Gamma_{\mathrm{GPRS}}(\boldsymbol{x},\boldsymbol{x}) &= \Gamma_k(\boldsymbol{x},\boldsymbol{x}) - \boldsymbol{\Gamma}_k^T(\boldsymbol{M}_k,\boldsymbol{x})\boldsymbol{\Gamma}_k^{-1}(\boldsymbol{M}_k,\boldsymbol{M}_k)\boldsymbol{\Gamma}_k(\boldsymbol{M}_k,\boldsymbol{x}).\end{aligned}$$

We define our response surface approximation to the posterior pdf as the mean

$$\hat{\pi}_{\mathrm{post}}(\boldsymbol{x}) = \mu_{\mathrm{GPRS}}(\boldsymbol{x}),$$

and our expression for the uncertainty in our approximation is the variance $\Gamma_{\mathrm{GPRS}}(\boldsymbol{x},\boldsymbol{x})$. We also define the nonnegative pdf approximation

$$\pi_{GPRS}(\boldsymbol{x}) = \begin{cases} \hat{\pi}_{\mathrm{post}}(\boldsymbol{x}) & \text{if } \hat{\pi}_{\mathrm{post}}(\boldsymbol{x}) \geq 0 \\ 0 & \text{if } \hat{\pi}_{\mathrm{post}}(\boldsymbol{x}) < 0 \end{cases}.$$

Since the posterior pdf is always nonnegative, this improves the approximation.

Note that, since we specified exact evaluations at each training point $\boldsymbol{x}_i$, the value of the mean of the posterior interpolation is $\mu_{\mathrm{GPRS}}(\boldsymbol{x}_i) = \pi_{\mathrm{post}}(\boldsymbol{x}_i)$, and the posterior variance is $\Gamma_{\mathrm{GPRS}}(\boldsymbol{x}_i,\boldsymbol{x}_i) = 0$. Furthermore, the first shape point chosen is the MAP point of the posterior pdf, and thus for a linear statistical inverse problem the response surface approximation at the first step of this algorithm can exactly recover the posterior pdf (a Gaussian whose covariance is the inverse of the Hessian and whose mean is the value of the posterior pdf at the MAP point: see Chapter 5 for a detailed exposition).

2.2.3 Greedy step initialization

We now return to the adaptive sampling algorithm described in §2.2.1 to discuss how to select an appropriate initial guess for the optimization problem in each greedy step. The first stage of the greedy algorithm solves the optimization problem

$$\boldsymbol{x}_{\text{new}} = \arg\max_{\boldsymbol{x}} \pi_{\text{post}}(\boldsymbol{x}).$$

This is none other than the deterministic inverse problem solving for the MAP point. Standard approaches for choosing initial guesses apply.

However, after finding the MAP point, the optimization problem for each subsequent greedy step is highly sensitive to the choice of initial guess. The result of the optimization algorithm will be a point at a local maximum of the difference between the posterior pdf and our Gaussian process response surface, but not necessarily at a point with the global maximum error. Due to the curse of dimensionality, we cannot cover the entire space, nor are randomly selected points likely to be useful. In a high dimensional parameter space, any pdf will be approximately zero almost everywhere. In these regions our response surface will naturally be an excellent approximation and an initial guess may prove useless. If the error between the posterior pdf and our approximation is too small, the optimization algorithm may not find a good search direction.

We note from Equation (2.1) that the Hessian is composed of two terms: the Hessian of the data misfit and the Hessian of the prior term. The Hessian can therefore be expressed as

$$\boldsymbol{H} = \boldsymbol{H}_{\text{misfit}} + \boldsymbol{\Gamma}_{\text{prior}}^{-1}. \tag{2.2}$$

We begin with a shape point $\boldsymbol{m}_k$, which places us in a region of interest. Directions around the shape point defined by the prior will already be well described by the GPRS. Instead, we search in the directions most informed by the data: the eigenvectors of the Hessian of the data misfit. The next question we must answer is how far to go in these directions. The GPRS approximates the Hessian locally, and thus if an initial guess is too close to the shape point the error in the approximation may be too small. If we travel too far from the shape point, then the posterior pdf and the GPRS will both converge to zero. We conclude that the eigenvectors should be scaled according to the local curvature and therefore we scale the eigenvectors by $1/\sqrt{\lambda}$, where λ is the associated eigenvalue.

Thus a set of potential good initial guesses given a shape point $\boldsymbol{m}_k$ is

$$M^0 = \{\boldsymbol{m}_k \pm \frac{1}{\sqrt{\lambda_j}}\boldsymbol{v}_j, \boldsymbol{m}_k \pm \frac{1}{n}\sum_j \frac{1}{\sqrt{\lambda_j}}\boldsymbol{v}_j\} \tag{2.3}$$

for a subset of the eigenpairs $(\lambda_j, \boldsymbol{v}_j)$. For each greedy step, we choose from the set of potential initial guesses generated from the available shape points using a criterion based on the relationship between our existing training points as described by the interpolation algorithm detailed earlier. As additional shape points are added to the GPRS, we expand this set of potential initial guesses using local Hessian information there.

Next we consider how to select one point from this set of potential initial guesses. We choose among them using an approximation by Christen and Sanso [18] to the Cohn criterion [20]. The initial guess $\boldsymbol{m}_{init}$ is defined as

$$\boldsymbol{m}_{init} = \max_{\boldsymbol{m}_j^0 \in M_0} J(\boldsymbol{m}_j^0) \tag{2.4}$$

where

$$J(\boldsymbol{m}_i^0) = \frac{\frac{1}{n^0}\sum_{j=1}^{n_0} \Gamma(\boldsymbol{m}_j^0, \boldsymbol{m}_i^0)^2 + \frac{1}{n_0}C_1}{C_2 + \sqrt{\sum_{j=1}^{n} \Gamma(\boldsymbol{m}_i^0, \boldsymbol{m}_j)^2}} \tag{2.5}$$

with $\boldsymbol{m}_j$ in the set of value points V and $\boldsymbol{m}_j^0$ in the set M^0 of potential initial guesses. C_1 and C_2 are constants independent of $\boldsymbol{m}_i^0$ and are defined by

$$C_1 = \sum_{i=1}^{n_0} \sum_{j=1}^{n_{values}} \Gamma(\boldsymbol{m}_i^0, \boldsymbol{m}_j)^2 \text{ and } C_2 = \max_{\boldsymbol{m}_i \in V} \sum_{j}^{n_{values}} |\Gamma(\boldsymbol{m}_i, \boldsymbol{m}_j)|.$$

This criteria selects for points with high predictive variance that are less correlated with points in the current set of value points but more correlated with other points in the set of potential initial guesses (and thus potentially provide the most new information for the response surface).

2.2.4 Gaussian mixture approximation

We now introduce a second approximation of $\pi_{\text{post}}(\boldsymbol{x})$ based on the shape points chosen through the greedy algorithm. Let the Gaussian mixture defined by the set of shape points be

$$\pi_{GM}(\boldsymbol{x}) \propto \sum w_j \exp\left(-\frac{1}{2}\|\boldsymbol{x} - \boldsymbol{m}_j\|_{\boldsymbol{H}_j}^2\right) \tag{2.6}$$

where

$$w_j = \pi_{\text{post}}(\boldsymbol{m_j})\sqrt{det(\boldsymbol{H}_j)}.$$

Unlike the GPRS, this is not an interpolation, nor does it include the information from the value points. It is thus less accurate. However, it is easy to draw samples from a Gaussian mixture, while we do not have an easy formula for drawing samples from the distribution defined by the normalization of the non-negative GPRS. Thus, we use this Gaussian mixture as the base of all three proposed Hessian-based MCMC algorithms.

2.3 Three MCMC algorithms employing Hessian-based response surfaces

We now consider the problem of calculating useful statistics of $\pi_{\text{post}}(\boldsymbol{x})$, such as the mean and variance, through MCMC. We described the construction of two approximations $\hat{\pi}_{\text{post}}(\boldsymbol{x})$ and $\pi_{GM}(\boldsymbol{x})$ to the posterior pdf $\pi_{\text{post}}(\boldsymbol{x})$ in §2.2. We employ these approximations in three MCMC algorithms with different tradeoffs between cost and accuracy. All three share the upfront cost of building the response surface.

We begin with the most accurate and most costly algorithm, namely Metropolis-Hastings sampling of $\pi_{\text{post}}(\boldsymbol{x})$ with $\pi_{GM}(\boldsymbol{x})$ as the proposal (see Algorithm 1 for step t of our proposed algorithm). This requires evaluation of the posterior pdf, which costs an additional PDE solve per sample over the upfront cost of building the Gaussian mixture.

Sampling the full posterior pdf directly is still quite expensive. For our second option, we substitute the response surface for the posterior pdf. We suggest Metropolis-Hastings sampling of $\pi_{GPRS}(\boldsymbol{x})$, with negative values set

Algorithm 1 Metropolis-Hastings sampling of $\pi_{\text{post}}(\boldsymbol{x})$ with $\pi_{GM}(\boldsymbol{x})$ as the proposal.

```
Given $\boldsymbol{x}_t$
Generate $\alpha$ from the uniform distribution
Generate sample $\boldsymbol{y}_t$ from $\pi_{GM}(\boldsymbol{y})$
Let
    $$\beta_t = \min\left(\frac{\pi_{\text{post}}(\boldsymbol{y}_t)\pi_{GM}(\boldsymbol{x}_t)}{\pi_{\text{post}}(\boldsymbol{x}_t)\pi_{GM}(\boldsymbol{y}_t)}, 1\right)$$
$\boldsymbol{x}_{t+1} = \begin{cases} \boldsymbol{x}_t & \text{if } \alpha < \beta_t \\ \boldsymbol{y}_t & \text{if } \alpha \geq \beta_t \end{cases}$
```

to zero, and the Gaussian mixture as a proposal density (see Algorithm 2). After constructing the response surface, no additional PDE solves are required, and thus the cost is much lower than the cost of Algorithm 1. This leaves the question of how closely the descriptive statistics from π_{GPRS} approximate those from $\pi_{\text{post}}(\boldsymbol{x})$, as we cannot calculate a global error. The question will be investigated numerically in Chapter 6.

Algorithm 2 Metropolis-Hastings sampling of $\hat{\pi}_{GPRS}(\boldsymbol{x}) = \max\{\hat{\pi}_{\text{post}}(\boldsymbol{x}), 0\}$ with proposal $\pi_{GM}(\boldsymbol{x})$.

```
Given $\boldsymbol{x}_t$
Generate $\alpha$ from the uniform distribution
Generate sample $\boldsymbol{y}_t$ from $\pi_{GM}(\boldsymbol{y})$

    $$\beta_t = \min\left(\frac{\hat{\pi}_{GPRS}(\boldsymbol{y}_t)\pi_{GM}(\boldsymbol{x}_t)}{\hat{\pi}_{GPRS}(\boldsymbol{x}_t)\pi_{GM}(\boldsymbol{y}_t)}, 1\right) \tag{2.7}$$
$\boldsymbol{x}_{t+1} = \begin{cases} \boldsymbol{x}_t & \text{if } \alpha < \beta_t \\ \boldsymbol{y}_t & \text{if } \alpha \geq \beta_t \end{cases}$
```

Our third option is a two stage delayed acceptance algorithm (Algorithm 3) to sample $\pi_{\text{post}}(\boldsymbol{x})$ which employs both $\pi_{GM}(\boldsymbol{x})$ and $\pi_{GPRS}(\boldsymbol{x})$. If the value of the response surface at a proposed point is negative, then we do not trust the response surface and hence in the accept/reject test we use the true posterior. Otherwise, if a point is accepted according to the response surface, we run a second accept/reject test with the posterior pdf. This is more accurate than sampling the response surface, and could reduce the cost compared to sampling the posterior pdf directly, since some points are rejected based on the response surface and thus do not cost an additional PDE solve.

Algorithm 3 Two stage delayed acceptance algorithm

```
Given $\boldsymbol{x}_t$
Generate $\alpha_0$ from the uniform distribution
Generate sample $\boldsymbol{y}_t$ from $\pi_{GM}(y)$
if $\pi_{GPRS}(\boldsymbol{y}_t) < 0$ then
    $\beta_t = \min\left(\frac{\pi_{\text{post}}(\boldsymbol{y}_t)\pi_{GM}(\boldsymbol{x}_t)}{\pi_{\text{post}}(\boldsymbol{x}_t)\pi_{GM}(\boldsymbol{y}_t)}, 1\right)$
    $\boldsymbol{x}_{t+1} = \begin{cases} \boldsymbol{x}_t \text{ if } \alpha_0 < \beta_t \\ \boldsymbol{y}_t \text{ if } \alpha_0 \geq \beta_t \end{cases}$
else
    $\beta_t^0 = \min\left(\frac{\pi_{GPRS}(\boldsymbol{y}_t)\pi_{GM}(\boldsymbol{x}_t)}{\pi_{GPRS}(\boldsymbol{x}_t)\pi_{GM}(\boldsymbol{y}_t)}, 1\right)$
    if $\alpha_0 < \beta_t^0$ then
        $\boldsymbol{x}_{t+1} = \boldsymbol{x}_t$
    else
        Generate $\alpha_1$ from the uniform distribution
        $\beta_t^1 = \min\left(\frac{\pi_{\text{post}}(\boldsymbol{y}_t)\pi_{GPRS}(\boldsymbol{x}^t)}{\pi_{\text{post}}(\boldsymbol{x}^t)\pi_{GPRS}(\boldsymbol{y}_t)}, 1\right)$
        $\boldsymbol{x}_{t+1} = \begin{cases} \boldsymbol{x}_t \text{ if } \alpha_1 < \beta_t^1 \\ \boldsymbol{y}_t \text{ if } \alpha_1 \geq \beta_t^1 \end{cases}$
    end if
end if
```

Chapter 3

Application to groundwater flow

In this chapter, we introduce the groundwater flow inverse problem (in §3.1). Then in §3.2 we develop expressions for the state, adjoint, gradient, incremental state, incremental adjoint, and Hessian. We conclude in §3.3 with a two parameter example in 1D to illustrate application of the algorithms in Chapter 2.

We target a large-scale nonlinear statistical inverse problem governed by single-phase fully saturated flow through an aquifer. The inverse problem is: given noisy measurements of the head at certain spatio-temporal locations, the pdf describing the noise, prior information in the form of a pdf of the hydraulic conductivity, and a groundwater flow model relating the head to the hydraulic conductivity, find the posterior pdf of the parameters describing the uncertainty in the hydraulic conductivity field[1].

3.1 Problem description

We first briefly describe some basic physical concepts in groundwater modeling [10]. The *piezometric head* is a measure of pressure, and is the sum

---

[1] The model uses log hydraulic conductivity to enforce positivity.

of the pressure head and the elevation head. The *pressure head* is the net work done by a unit weight of incompressible fluid against the pressure difference existing along its flow. The *elevation head* is the potential energy per unit weight of fluid. The *log hydraulic conductivity* $\gamma = \log(kg/\nu)$ includes the effects of the permeability k of the rock matrix, the dynamic viscosity ν, and the gravity g and is thus a property of both the rock and the fluid. For two dimensional models (when flow is essentially horizontal, and thus the Dupuit approximation is appropriate), the *transmissivity* $T = \bar{K}b$ is obtained from $\bar{K}$, hydraulic conductivity averaged over height (assuming a homogeneous fluid), and the vertical thickness b of the aquifer. We refer the reader to [10], page 136, for typical values of hydraulic conductivity. The *storativity* S of the aquifer is the change in the volume of water due to change in the piezometric head. The storativity is the outcome of elastic properties of the medium and the water. Usually, it is assumed to be time independent, but it may vary spatially.

For our target application, the forward problem is governed by single-phase fully saturated flow through an aquifer:

$$
\begin{aligned}
S\frac{\partial h}{\partial t} - \nabla \cdot (e^{\gamma}\nabla h) &= f, && \Omega \times (0, T), \qquad (3.1)\\
h &= h_0, && \Omega \times \{t = 0\},\\
e^{\gamma}\nabla h \cdot \mathbf{n} &= h_N, && \Gamma_N \times (0, T),\\
h &= h_D, && \Gamma_D \times (0, T),
\end{aligned}
$$

where S is the specific storativity, f describes sources and sinks, γ is the log hydraulic conductivity, and h is the piezometric head.

As described in Chapter 2, the Bayesian formulation of a nonlinear inverse problem with Gaussian noise and prior is related to an appropriately weighted least squares minimization problem. We choose to define our noise and prior pdf's by discretizing the infinite-dimensional functional

$$\frac{1}{2}\int_0^T \int_\Omega (h - h_{\text{obs}} - \bar{e})^2 b_{\text{noise}}\, d\boldsymbol{x}\, dt + \frac{1}{2}\int_\Omega (A(\gamma - \bar{\gamma}_{\text{prior}}))^2 d\boldsymbol{x}, \tag{3.2}$$

in which $h(x,t)$ satisfies the single-phase flow equation 3.1. The operator A is a second order elliptic differential operator which in weak form is defined as: for $s \in L^2(\Omega)$ the solution $\gamma = A^{-1}s$ satisfies

$$\int_\Omega \beta_{0,\text{prior}}\gamma\psi + \beta_{1,\text{prior}}\nabla\gamma \cdot \nabla\psi\, d\boldsymbol{x} = \int_\Omega s\psi\, d\boldsymbol{x} \text{ for all } \psi \in H^1(\Omega)$$

where $\beta_{0,\text{prior}}$ and $\beta_{1,\text{prior}}$ are positive constants. This is an appropriate choice to ensure well-posedness of the infinite dimensional Bayesian formulation [51] for a 2D problem. A more detailed discussion of this choice of prior is available in [16]. This operator also provides a simple and fast square root operator and allows a straightforward discretization. Discussion of an appropriate lower order prior for one-dimensional inversion fields can be found, for example, in [49].

We define the observation operator as

$$b(x,t) = \sum_j \delta(x - x_j).$$

Then discretization of expression (3.2) is equivalent to the choice of finite-dimensional Gaussian pdf's for the prior and noise

$$\pi_{\text{prior}}(\boldsymbol{\gamma}) = \mathcal{N}(\bar{\gamma}_{\text{prior}}, \boldsymbol{\Gamma}_{\text{prior}})$$
$$\pi_{\text{noise}}(\boldsymbol{e}) = \mathcal{N}(0, \boldsymbol{\Gamma}_{\text{noise}})$$

where the prior and noise covariance matrices are defined as

$$\boldsymbol{\Gamma}_{\text{prior}} = \boldsymbol{K}\boldsymbol{M}^{-1}\boldsymbol{K},$$
$$\boldsymbol{\Gamma}_{\text{noise}} = \beta_{\text{noise}}\boldsymbol{I}.$$

Here $\boldsymbol{M}$ and $\boldsymbol{K}$ are the mass and stiffness matrices respectively, whose components are given by

$$\begin{aligned}\boldsymbol{K}_{ij} &= \int_{\Omega} (\beta_{0,\text{prior}}\psi_i\psi_j + \beta_{1,\text{prior}}\nabla\psi_i \cdot \nabla\psi_j)\, dx \\ \boldsymbol{M}_{ij} &= \int_{\Omega} \psi_i\psi_j dx,\end{aligned}$$

with appropriately chosen finite element basis functions $\{\psi_i\}_{i=1}^{n}$.

Thus, we address the statistical inverse problem for the posterior pdf of the hydraulic conductivity in finite dimensions, after discretizing the forward model to create a "parameter-to-observable" map. From Bayes' theorem, an additive Gaussian noise pdf and Gaussian prior pdf, results in an expression for the (non-Gaussian) posterior pdf,

$$\pi_{\text{post}}(\boldsymbol{\gamma}) \propto \exp\left(-\frac{1}{2}\|\boldsymbol{h}_{\text{obs}} - \boldsymbol{h}(\boldsymbol{\gamma})\|^2_{\boldsymbol{\Gamma}^{-1}_{\text{noise}}} - \frac{1}{2}\|\boldsymbol{\gamma} - \bar{\boldsymbol{\gamma}}_{\text{prior}}\|^2_{\boldsymbol{\Gamma}^{-1}_{\text{prior}}}\right), \tag{3.3}$$

whose evaluation requires application of the parameter-to-observable map $\boldsymbol{h}(\gamma)$ (which implies solution of the forward problem (3.1) for a given γ).

In later numerical results for our statistical inversion methodology, we work with the steady-state version of the forward equation. The steady-state groundwater flow model is implemented in C++ and MPI using the deal II [9, 8] and PETSc [5, 6, 7] libraries with linear finite elements, while parallel implementation of the adaptive Hessian-informed Bayesian interpolation algorithm is done in C++ and MPI using PETSc [5, 6, 7], with SLEPc [33] for the low rank Hessian approximation and the GNU Scientific Library (GSL) for random number generation needed by the MCMC algorithms.

3.2 Derivation of reduced Hessian

We now develop expressions for the forward, adjoint, incremental forward, incremental adjoint, reduced gradient, and reduced Hessian through the method of Lagrange multipliers.

The Bayesian formulation of a statistical inverse problem with Gaussian noise and prior is related to an appropriately-weighted nonlinear least squares minimization problem. We define the infinite dimensional functional $J(h, \gamma)$ as the objective function

$$J(h, \gamma) = \frac{1}{2} \int_0^T \int_\Omega (h - h_{\text{obs}} - \bar{e})^2 b_{\text{noise}} \, d\boldsymbol{x} \, dt + R(\gamma, \bar{\gamma}_{\text{prior}}),$$

where $R(\gamma, \bar{\gamma}_{\text{prior}})$ is determined by prior information.

The associated deterministic PDE constrained optimization problem is

$$\min_{\gamma} J(h, \gamma),$$

in which $h(x, t)$ satisfies the single-phase flow equation 3.1. The placement of sensors to measure observations h_{obs} of the head, and the noise in their measurements, is described by b_{noise}. The error $\bar{e}$ is the mean of the noise pdf. β_{prior} and $\bar{\gamma}_{\text{prior}}$ are parameters that determine the strength and shape of the regularization term.

For both the deterministic and Bayesian framework, we approach the problem using the Lagrange method for equality constrained optimization problems. From this we derive optimality conditions, discretized state and adjoint equations, the KKT system, and finally the reduced Hessian. We introduce adjoint variables q and r corresponding to the PDE constraint and the initial condition restraint. Then we define the Lagrangian functional as

$$\begin{aligned}\mathcal{L}(h, q, r, \gamma) = &\frac{1}{2}\int_0^T \int_\Omega (h - h_{\text{obs}} - \bar{e})^2 b_{\text{noise}} \, d\boldsymbol{x} \, dt + R(\gamma, \bar{\gamma}_{\text{prior}}) \\ &+ \int_0^T \int_\Omega \left(Sq\frac{\partial h}{\partial t} - \nabla \cdot (e^\gamma \nabla h) q - fq \right) d\boldsymbol{x} \, dt \\ &+ \int_\Omega r(h(x, 0) - h_0) d\boldsymbol{x}\end{aligned}$$

Integrating by parts, we obtain the expression

$$\begin{aligned}\mathcal{L}(h, q, r, \gamma) = &\frac{1}{2}\int_0^T \int_\Omega (h - h_{\text{obs}} - \bar{e})^2 b_{\text{noise}} \, d\boldsymbol{x} \, dt + R(\gamma, \bar{\gamma}_{\text{prior}}) \\ &+ \int_0^T \int_\Omega \left(Sq\frac{\partial h}{\partial t} + e^\gamma \nabla h \cdot \nabla q - fq \right) d\boldsymbol{x} \, dt \\ &- \int_0^T \int_\Gamma e^\gamma \nabla h \cdot \boldsymbol{n} q ds dt + \int_\Omega r(h(x, 0) - h_0) d\boldsymbol{x}.\end{aligned}$$

and the function spaces are defined as

$V = \{v \in H^1(\Omega) : v|_{\Gamma_D} = 0\}$

state variable: $h \in L^2(0, T; V + \hat{h}_D)$ where $\hat{h}_D$ is the lift function for h_D

adjoint variables: $q \in L^2(0, T; V)$, $r \in L^2(\Omega)$

control variable: $\gamma \in L^2(\Omega)$

forcing function: $f \in L^2(0, T; L^2(\Omega))$

initial condition: $h_0 \in L^2(\Omega)$

Neumann condition: $h_N \in H^{-1/2}(\Gamma_N)$

Dirichlet condition: $h_D \in H^{1/2}(\Gamma_D)$.

Recall that the Neumann boundary condition is $e^{\gamma}\nabla h \cdot \boldsymbol{n} = h_N$ on Γ_N. Thus,

$$\begin{aligned}\mathcal{L}(h, q, r, \gamma) = \frac{1}{2}\int_0^T \int_\Omega (h - h_{\text{obs}} - \bar{e})^2 b_{\text{noise}}\, d\boldsymbol{x}\, dt + R(\gamma, \bar{\gamma}_{\text{prior}}) \\ + \int_0^T \int_\Omega \left(Sq\frac{\partial h}{\partial t} + e^{\gamma}\nabla h \cdot \nabla q - fq \right) d\boldsymbol{x}\, dt \\ - \int_0^T \int_{\Gamma_N} h_N q ds dt + \int_\Omega r(h(x, 0) - h_0) d\boldsymbol{x}\end{aligned}$$

3.2.1 Derivation of the reduced gradient

We are interested in evaluating the reduced gradient, $dJ/d\gamma$. From the gradient of the Lagrangian functional we obtain a system of equations based on its variations with respect to h, r, and q:

$$\begin{aligned} &\mathcal{L}_q = 0 && \text{state equation} \\ &\mathcal{L}_r = 0 && \text{state initial condition} \\ &\mathcal{L}_h = 0 && \text{adjoint equation} \end{aligned}$$

The reduced gradient $dJ/d\gamma$ is obtained by evaluating the expression for $\delta_\gamma \mathcal{L}$, the variation of the Lagrangian function with respect to γ, with the solutions h and q from the state and adjoint equations.

3.2.1.1 Derivation and discretization of the state equation

We derive the state equation by taking the variation of the objective function $\mathcal{L}$ with respect to the adjoint variable q as follows:

$$\begin{aligned} \mathcal{L}_q(\hat{q}) = &\int_0^T \int_\Omega \left(S \frac{\partial h}{\partial t} \hat{q} + e^\gamma \nabla h \cdot \nabla \hat{q} - f \hat{q} \right) d\boldsymbol{x} dt \\ &- \int_0^T \int_{\Gamma_N} h_N \hat{q} ds\, dt, \qquad \forall \hat{q} \in L^2(0,T;V). \end{aligned}$$

We derive initial conditions for the state equation from variations of $\mathcal{L}$ with respect to r:

$$\mathcal{L}_r(\hat{r}) = \int_\Omega \hat{r}(h(x,0) - h_0(x)) d\boldsymbol{x} = 0, \qquad \forall\, \hat{r} \in L^2(\Omega).$$

Thus, the weak form of the state equation is

$$\begin{aligned} &\int_0^T \int_\Omega \left(S \frac{\partial h}{\partial t} \hat{q} + e^\gamma \nabla h \cdot \nabla \hat{q} - f \hat{q} \right) d\boldsymbol{x}\, dt \\ &- \int_0^T \int_{\Gamma_N} h_N \hat{q} ds\, dt = 0, \qquad \forall \hat{q} \in L^2(0,T;V) \\ &\int_\Omega \hat{r}(h(x,0) - h_0(x)) d\boldsymbol{x} = 0, \qquad \forall\, \hat{r} \in L_2(\Omega) \end{aligned}$$

The strong form of the state equation is

$$\begin{aligned} S\frac{\partial h}{\partial t} - \nabla \cdot (e^{\gamma} \nabla h) &= f, && \Omega \times (0,T) \\ h(x,0) &= h_0(x), && \Omega \times \{t=0\} \\ e^{\gamma}\nabla h \cdot \boldsymbol{n} &= h_N, && \Gamma_N \times (0,T), \\ h(x,t) &= h_D && \Gamma_D \times (0,T) \end{aligned}$$

We discretize in space with Galerkin finite elements: $h = \sum_j h_j \psi_j(x)$. Substituting into the weak form, we obtain

$$\boldsymbol{M}_{\text{mass}} \frac{\partial \boldsymbol{h}}{\partial t} = -\boldsymbol{K}_{\text{stiff}}(\boldsymbol{\gamma})\boldsymbol{h} + \boldsymbol{F}_{\text{state}},$$

where

$$\begin{aligned} (\boldsymbol{M}_{\text{mass}})_{ij} &= \int_{\Omega} S\psi_i\psi_j d\boldsymbol{x}, \\ (\boldsymbol{K}_{\text{state}})_{ij} &= \int_{\Omega} e^{\gamma}\nabla\psi_i \nabla\psi_j d\boldsymbol{x} \\ (\boldsymbol{F}_{\text{state}})_i &= \int_{\Omega} f\psi_i d\boldsymbol{x} + \int_{\Gamma_N} h_N \psi_i ds. \end{aligned}$$

For the steady-state case, $\partial \boldsymbol{h}/\partial t = 0$, and thus this expression becomes

$$\boldsymbol{K}_{\text{state}}(\boldsymbol{\gamma})\boldsymbol{h} = \boldsymbol{F}_{\text{state}}.$$

In the time-dependent case, we use the second-order scheme Crank-Nicolson for time discretization:

$$\boldsymbol{M}_{\text{mass}} \left(\frac{\boldsymbol{h}^{t+1} - \boldsymbol{h}^{t}}{\delta t} \right) = \frac{1}{2}(-\boldsymbol{K}_{\text{state}}(\boldsymbol{\gamma})\boldsymbol{h}^{t+1} - \boldsymbol{K}_{\text{state}}(\boldsymbol{\gamma})\boldsymbol{h}^{t} + \boldsymbol{F}^{t+1}_{\text{state}} + \boldsymbol{F}^{t}_{\text{state}}).$$

Rearranging, we obtain a final equation

$$\begin{aligned}\left(\boldsymbol{M}_{\text{mass}} + \frac{\delta t}{2}\boldsymbol{K}_{\text{state}}(\boldsymbol{\gamma})\right)\boldsymbol{h}^{t+1} = \left(\boldsymbol{M}_{\text{mass}} - \frac{\delta t}{2}\boldsymbol{K}_{\text{state}}(\boldsymbol{\gamma})\right)\boldsymbol{h}^{t} \\ + \frac{\delta t}{2}(\boldsymbol{F}^{t}_{\text{state}} + \boldsymbol{F}^{t+1}_{\text{state}}).\end{aligned}$$

3.2.1.2 Derivation and discretization of the adjoint equation

We next derive the adjoint equation from the variation of $\mathcal{L}$ with respect to h, and integrate by parts in time:

$$\begin{aligned}\mathcal{L}_h(\hat{h}) = &\int_0^T \int_\Omega b_{\text{noise}}\hat{h}(h - h_{\text{obs}} - \bar{e})\, d\boldsymbol{x}\, dt \\ &+ \int_0^T \int_\Omega (-S\frac{\partial q}{\partial t}\hat{h} + e^{\gamma}\nabla\hat{h}\cdot\nabla q) d\boldsymbol{x}\, dt \\ &+ \int_\Omega (Sq(x,T)\hat{h}(x,T) - Sq(x,0)\hat{h}(x,0)) d\boldsymbol{x} \\ &+ \int_\Omega r\hat{h}(x,0) d\boldsymbol{x} = 0, \qquad \forall\, \hat{h} \in L^2(0,T;V).\end{aligned}$$

Integrating by parts in space,

$$\begin{aligned}\mathcal{L}_h(\hat{h}) = &\int_0^T \int_\Omega b_{\text{noise}}\hat{h}(h - h_{\text{obs}} - \bar{e})\, d\boldsymbol{x}\, dt \\ &+ \int_0^T \int_\Omega (-S\frac{\partial q}{\partial t} - \nabla\cdot(e^{\gamma}\nabla q))\hat{h} d\boldsymbol{x}\, dt \\ &+ \int_\Omega (Sq(x,T)\hat{h}(x,T) - Sq(x,0)\hat{h}(x,0)) d\boldsymbol{x} \\ &+ \int_0^T \int_\Gamma e^{\gamma}\nabla q\cdot\boldsymbol{n}\hat{h} ds\, dt \\ &+ \int_\Omega r\hat{h}(x,0) d\boldsymbol{x} = 0, \qquad \forall\, \hat{h} \in L^2(0,T;V).\end{aligned}$$

We then separate terms into a series of expressions based on associated test functions:

$$\begin{aligned}
&\int_0^T \int_\Omega b_{\text{noise}}\hat{h}(h - h_{\text{obs}} - \bar{e})\, d\boldsymbol{x}\, dt + \int_0^T \int_\Omega (-S\frac{\partial q}{\partial t} - \nabla \cdot (e^\gamma \nabla q))\hat{h} d\boldsymbol{x}\, dt = 0 \\
&\int_0^T \int_{\Gamma_N} e^\gamma \nabla q \cdot \boldsymbol{n}\hat{h} ds\, dt = 0 \\
&\int_\Omega (Sq(x,T)\hat{h}(x,T)dx = 0 \\
&- \int_\Omega Sq(x,0)\hat{h}(x,0))d\boldsymbol{x} + \int_\Omega r\hat{h}(x,0)d\boldsymbol{x} = 0, \qquad \forall\, \hat{h} \in L^2(0,T;V).
\end{aligned}$$

From this we discover that the Lagrange multipliers q and r are related by the expression $r(x) = Sq(x,0)$ in the sense of the trace. We also obtain the final time condition $q(x,T) = 0$ and the Neumann boundary condition $e^\gamma \nabla q \cdot \boldsymbol{n} = 0$ on Γ_N.

The strong form of the adjoint equation is

$$\begin{aligned}
S\frac{\partial q}{\partial t} + \nabla \cdot (e^\gamma \nabla q) &= (h - h_{\text{obs}} - \bar{e})b_{\text{noise}}, && \Omega \times (0,T) \\
q(T) &= 0, && \Omega \times \{t = T\}, \\
e^\gamma \nabla q \cdot \boldsymbol{n} &= 0, && \Gamma_N \times (0,T) \\
q &= 0 && \Gamma_D \times (0,T).
\end{aligned}$$

The weak form of the adjoint equation is given by

$$\begin{aligned}
&\int_0^T \int_\Omega b_{\text{noise}}\hat{h}(h - h_{\text{obs}} - \bar{e})\, d\boldsymbol{x}\, dt + \int_0^T \int_\Omega (-S\frac{\partial q}{\partial t}\hat{h} + e^\gamma \nabla \hat{h} \cdot \nabla q) d\boldsymbol{x}\, dt \\
&+ \int_\Omega Sq(x,T)\hat{h}(x,T)d\boldsymbol{x} - \int_0^T \int_{\Gamma_D} \hat{h} e^\gamma \nabla q \cdot \boldsymbol{n} ds\, dt = 0, \qquad \forall\, \hat{h} \in L^2(0,T;V).
\end{aligned}$$

We use a Galerkin finite element discretization: $q = \sum_k h_k \psi_k$. Substituting into the weak form, we obtain

$$\boldsymbol{M}_{\text{mass}} \frac{\partial \boldsymbol{q}}{\partial t} = \boldsymbol{K}_{\text{adjoint}}(\boldsymbol{\gamma})\boldsymbol{q} + \boldsymbol{B}\boldsymbol{h} - \boldsymbol{F}_{\text{adjoint}},$$

where

$$\begin{aligned}(\boldsymbol{M}_{\text{mass}})_{ij} = \int_\Omega S\psi_i\psi_j d\boldsymbol{x}, (\boldsymbol{F}_{\text{adjoint}})_i &= \int_\Omega (h_{\text{obs}} + \bar{e}) b_{\text{noise}} \psi_i \, d\boldsymbol{x} \\ \boldsymbol{B}_{ij} &= \int_\Omega b_{\text{noise}} \psi_i \psi_j \, d\boldsymbol{x} \\ \boldsymbol{K}_{\text{adjoint}} &= \int_\Omega e^{\gamma} \nabla\psi_i \cdot \nabla\psi_j d\boldsymbol{x}\end{aligned}$$

For steady-state, the expression becomes

$$\boldsymbol{K}_{\text{adjoint}}(\boldsymbol{\gamma})\boldsymbol{q} = -(\boldsymbol{B}\boldsymbol{h} - \boldsymbol{F}_{\text{adjoint}}).$$

For the time-dependent case, we discretize in time with Crank-Nicolson as follows:

$$\begin{aligned}\boldsymbol{M}_{\text{mass}} \left(\frac{\boldsymbol{q}^{t+1} - \boldsymbol{q}^t}{\delta t} \right) &= \frac{1}{2} \left(\boldsymbol{K}_{\text{adjoint}}(\boldsymbol{\gamma})\boldsymbol{q}^{t+1} + \boldsymbol{K}_{\text{adjoint}}(\boldsymbol{\gamma})\boldsymbol{q}^t \right) \\ &+ \frac{1}{2} \left(\boldsymbol{B}^{t+1}\boldsymbol{h}^{t+1} + \boldsymbol{B}^t\boldsymbol{h}^t - \boldsymbol{F}^{t+1}_{\text{adjoint}} - \boldsymbol{F}^t_{\text{adjoint}} \right)\end{aligned}$$

Rearranging, we obtain a final expression

$$\begin{aligned}\left(\boldsymbol{M}_{\text{mass}} + \frac{\delta t}{2} \boldsymbol{K}_{\text{adjoint}}(\boldsymbol{\gamma}) \right) \boldsymbol{q}^t = &- \left(\boldsymbol{M}_{\text{mass}} - \frac{\delta t}{2} \boldsymbol{K}_{\text{adjoint}}(\boldsymbol{\gamma}) \right) \boldsymbol{q}^{t+1} \\ &- \frac{\delta t}{2} \left(\boldsymbol{B}^t\boldsymbol{h}^t + \boldsymbol{B}^{t+1}\boldsymbol{h}^{t+1} - \boldsymbol{F}^t_{\text{adjoint}} - \boldsymbol{F}^{t+1}_{\text{adjoint}} \right).\end{aligned}$$

3.2.1.3 Derivation of the reduced gradient

We now substitute the relationships between Lagrange multipliers q and r into the Lagrangian functional. The expression for the Lagrangian functional becomes

$$\begin{aligned}\mathcal{L}(h, q, \gamma) = &\frac{1}{2}\int_0^T \int_\Omega (h - h_{\text{obs}} - \bar{e})^2 b_{\text{noise}}\, d\boldsymbol{x}\, dt + R(\gamma, \bar{\gamma}_{\text{prior}}) \\ &+ \int_0^T \int_\Omega \left(Sq\frac{\partial h}{\partial t} + e^\gamma \nabla h \cdot \nabla q - fq \right) d\boldsymbol{x}\, dt \\ &+ \int_\Omega Sq(x, 0)(h(x, 0) - h_0) d\boldsymbol{x} - \int_0^T \int_{\Gamma_N} h_N q\, ds dt.\end{aligned}$$

We derive the variation of the Lagrangian function with respect to γ:

$$\mathcal{L}_\gamma(\hat{\gamma}) = \int_0^T \int_\Omega \hat{\gamma} e^\gamma \nabla h \cdot \nabla q\, d\boldsymbol{x} dt + R_\gamma(\gamma, \hat{\gamma}).$$

The discretized reduced gradient for a given $\bar{h}$ and $\bar{q}$ satisfying the state and adjoint equations is therefore

$$\boldsymbol{C}(\boldsymbol{\gamma}, \boldsymbol{h}, \boldsymbol{q}) + \boldsymbol{R}(\gamma, \hat{\gamma})$$

where

$$\boldsymbol{C}(\boldsymbol{\gamma}, \boldsymbol{h}, \boldsymbol{q})_i = \int_0^T \int_\Omega \psi_i e^\gamma \nabla h \cdot \nabla q\, d\boldsymbol{x} dt.$$

3.2.2 Derivation of the reduced Hessian

The KKT system with the full Hessian is

$$\begin{pmatrix} \mathcal{L}_{hh} & \mathcal{L}_{hq} & \mathcal{L}_{h\gamma} \\ \mathcal{L}_{qh} & \mathcal{L}_{qq} & \mathcal{L}_{q\gamma} \\ \mathcal{L}_{\gamma h} & \mathcal{L}_{\gamma q} & \mathcal{L}_{\gamma\gamma} \end{pmatrix} \begin{pmatrix} \tilde{h} \\ \tilde{q} \\ \tilde{\gamma} \end{pmatrix} = - \begin{pmatrix} \mathcal{L}_h \\ \mathcal{L}_q \\ \mathcal{L}_\gamma \end{pmatrix},$$

3.2.2.1 Incremental state equation

Thus the incremental state equation for $\tilde{h} \in L^2(0, T; V)$ is

$$\mathcal{L}_q(h) + \mathcal{L}_{qh}(\tilde{h}) + \mathcal{L}_{qq}(\tilde{q}) + \mathcal{L}_{q\gamma}(\tilde{\gamma}) = 0, \qquad \forall \hat{h} \in L^2(0, T; V),$$

where h is the solution to the state equation and therefore $\mathcal{L}_q(h) = 0$. The incremental state equation is built from the following variations:

$$\begin{aligned}
&\mathcal{L}_{qq}(\tilde{q}) = 0 \\
&\mathcal{L}_{qh}(\tilde{h}) = \int_0^T \int_\Omega \left(S\hat{q}\frac{\partial \tilde{h}}{\partial t} + e^\gamma \nabla \tilde{h} \cdot \nabla \hat{q} \right) d\boldsymbol{x}\, dt + \int_\Omega S\hat{q}(x, 0)\tilde{h}(x, 0)\, d\boldsymbol{x} \\
&\mathcal{L}_{q\gamma}(\tilde{\gamma}) = \int_0^T \int_\Omega \tilde{\gamma} e^\gamma \nabla h \cdot \nabla \hat{q}\, d\boldsymbol{x}\, dt
\end{aligned}$$

The weak form is thus

$$\begin{aligned}
&\int_0^T \int_\Omega \left(S\hat{q}\frac{\partial \tilde{h}}{\partial t} + e^\gamma \nabla \tilde{h} \cdot \nabla \hat{q} + \tilde{\gamma} e^\gamma \nabla h \cdot \nabla \hat{q} \right) d\boldsymbol{x}\, dt \\
&+ \int_\Omega S\hat{q}(x, 0)\tilde{h}(x, 0)\, d\boldsymbol{x} \qquad \forall \hat{q} \in L^2(0, T; V).
\end{aligned}$$

Discretizing in space,

$$\boldsymbol{M}_{\text{mass}} \frac{\partial \tilde{\boldsymbol{h}}}{\partial t} + \boldsymbol{K}_{\text{incremental}}(\boldsymbol{\gamma})\tilde{\boldsymbol{h}} + \boldsymbol{C}_{\text{inc state}}(\boldsymbol{h}, \boldsymbol{\gamma})\tilde{\boldsymbol{\gamma}} = 0,$$

where

$$\begin{aligned}
\boldsymbol{C}_{\text{inc state}}(\boldsymbol{h}, \boldsymbol{\gamma}) &= \int_\Omega \psi_j e^\gamma \nabla h \cdot \nabla \psi_i d\boldsymbol{x} \\
\boldsymbol{K}_{\text{incremental}}(\boldsymbol{\gamma}) &= \int_\Omega e^\gamma \nabla \psi_i \cdot \nabla \psi_j d\boldsymbol{x}
\end{aligned}$$

The steady-state incremental state equation is

$$\boldsymbol{K}_{\text{incremental}}(\boldsymbol{\gamma})\tilde{\boldsymbol{h}} = -\boldsymbol{C}_{\text{inc state}}(\boldsymbol{h}, \boldsymbol{\gamma})\tilde{\boldsymbol{\gamma}}.$$

Integrating the weak form by parts, we obtain

$$\begin{aligned}
0 = &\int_0^T \int_\Omega \left(S\hat{q}\frac{\partial \tilde{h}}{\partial t} - \left(\nabla \cdot e^\gamma \nabla \tilde{h}\right) \hat{q} - \left(\nabla \cdot \tilde{\gamma} e^\gamma \nabla h\right) \hat{q} \right) d\boldsymbol{x}\, dt \\
&+ \int_\Omega S\hat{q}(x,0)\tilde{h}(x,0)\, d\boldsymbol{x} \\
&+ \int_0^T \int_\Gamma e^\gamma \nabla \tilde{h} \cdot \boldsymbol{n}\hat{q} ds\, dt + \int_0^T \int_\Gamma \tilde{\gamma} e^\gamma \nabla h \cdot \boldsymbol{n}\hat{q} ds\, dt \\
= &\int_0^T \int_\Omega \left(S\hat{q}\frac{\partial \tilde{h}}{\partial t} - \left(\nabla \cdot e^\gamma \nabla \tilde{h}\right) \hat{q} - \left(\nabla \cdot \tilde{\gamma} e^\gamma \nabla h\right) \hat{q} \right) d\boldsymbol{x}\, dt \\
&+ \int_\Omega S\hat{q}(x,0)\tilde{h}(x,0)\, d\boldsymbol{x} \\
&+ \int_0^T \int_{\Gamma_N} e^\gamma \nabla \tilde{h} \cdot \boldsymbol{n}\hat{q} ds\, dt + \int_0^T \int_{\Gamma_N} \tilde{\gamma} e^\gamma \nabla h \cdot \boldsymbol{n}\hat{q} ds\, dt \\
&\quad \forall \hat{q} \in L^2(0,T;V).
\end{aligned}$$

The strong form of the incremental state equation is thus

$$\begin{aligned}
S\frac{\partial \tilde{h}}{\partial t} - \nabla \cdot (e^\gamma \nabla \tilde{h}) &= \nabla \cdot (\tilde{\gamma} e^\gamma \nabla h) && \Omega \times (0,T) \\
S\tilde{h}(x,0) &= 0, && \Omega \times \{t=0\} \\
\tilde{h} &= 0 && \Gamma_D \times (0,T) \\
e^\gamma \nabla \tilde{h} \cdot \boldsymbol{n} &= -\tilde{\gamma} h_N && \Gamma_N \times (0,T)
\end{aligned}$$

3.2.2.2 Incremental adjoint equation

The incremental adjoint equation for $\tilde{q} \in L^2(0, T; V)$ is

$$\mathcal{L}_h(\bar{q}) + \mathcal{L}_{hq}(\tilde{h}) + \mathcal{L}_{hh}(\tilde{q}) + \mathcal{L}_{h\gamma}(\tilde{\gamma}) = 0, \qquad \forall \hat{q} \in L^2(0, T; V),$$

where q is the solution to the adjoint equation and therefore $\mathcal{L}_h(q) = 0$. This expression is built from the following variations:

$$\begin{aligned}
\mathcal{L}_{hh}(\tilde{h}) &= \int_0^T \int_\Omega b_{\text{noise}} \hat{h} \tilde{h}\, d\boldsymbol{x} dt \\
\mathcal{L}_{hq}(\tilde{q}) &= \int_0^T \int_\Omega \left(-S \frac{\partial \tilde{q}}{\partial t} \hat{h} + e^\gamma \nabla \hat{h} \cdot \nabla \tilde{q} \right) d\boldsymbol{x}\, dt + \int_\Omega S \tilde{q}(x, T) \hat{h}(x, T) d\boldsymbol{x} \\
\mathcal{L}_{h\gamma}(\tilde{\gamma}) &= \int_0^T \int_\Omega \tilde{\gamma} e^\gamma \nabla \hat{h} \cdot \nabla q d\boldsymbol{x}\, dt
\end{aligned}$$

Combining the variations, we obtain the expression

$$\begin{aligned}
&\int_0^T \int_\Omega b_{\text{noise}} \hat{h} \tilde{h}\, d\boldsymbol{x} dt + \int_0^T \int_\Omega \left(-S \frac{\partial \tilde{q}}{\partial t} \hat{h} + e^\gamma \nabla \hat{h} \cdot \nabla \tilde{q} \right) d\boldsymbol{x}\, dt \\
&+ \int_\Omega S \tilde{q}(x, T) \hat{h}(x, T) dx + \int_0^T \int_\Omega \tilde{\gamma} e^\gamma \nabla \hat{h} \cdot \nabla q d\boldsymbol{x}\, dt = 0 \qquad \forall \hat{h} \in L^2(0, T; V)
\end{aligned}$$

Discretizing in space,

$$-\boldsymbol{M}_{\text{mass}} \frac{\partial \tilde{\boldsymbol{q}}}{\partial t} + \boldsymbol{K}_{\text{incremental}} \tilde{\boldsymbol{q}} + \boldsymbol{C}_{\text{inc adjoint}}(\boldsymbol{q}, \boldsymbol{\gamma}) \tilde{\boldsymbol{\gamma}} = -\boldsymbol{B} \tilde{\boldsymbol{h}},$$

where

$$\begin{aligned}
\boldsymbol{B}_{ij} &= \int_\Omega b_{\text{noise}} \psi_i \psi_j d\boldsymbol{x}, \\
\boldsymbol{C}_{\text{inc adjoint}}(\boldsymbol{q}, \boldsymbol{\gamma}) &= \int_\Omega \psi_j e^\gamma \nabla q \cdot \nabla \psi_i d\boldsymbol{x} \\
\boldsymbol{K}_{\text{incremental}}(\boldsymbol{\gamma}) &= \int_\Omega e^\gamma \nabla \psi_i \cdot \nabla \psi_j d\boldsymbol{x}
\end{aligned}$$

Therefore, the discretized steady-state incremental adjoint equation is

$$\boldsymbol{K}_{\text{incremental}}\tilde{\boldsymbol{q}} = -\boldsymbol{C}_{\text{inc adjoint}}(\boldsymbol{q},\boldsymbol{\gamma})\tilde{\boldsymbol{\gamma}} - \boldsymbol{B}\tilde{\boldsymbol{h}}.$$

Integrating the weak form by parts,

$$\begin{aligned}
0 &= \int_0^T \int_\Omega \hat{h}\tilde{h}b_{noise} d\boldsymbol{x}\,dt - \int_0^T \int_\Omega \left(S\frac{\partial \tilde{q}}{\partial t} + \nabla\cdot e^{\gamma}\nabla\tilde{q} + \nabla\cdot\tilde{\gamma}e^{\gamma}\nabla q\right)\hat{h} d\boldsymbol{x}\,dt \\
&+ \int_\Omega S\tilde{q}(x,T)\hat{h}(x,T)d\boldsymbol{x} + \int_0^T \int_\Gamma e^{\gamma}\nabla\tilde{q}\cdot\boldsymbol{n}\hat{h} ds\,dt \\
&+ \int_0^T \int_\Gamma \tilde{\gamma}e^{\gamma}\nabla q\cdot\boldsymbol{n}\hat{h} ds\,dt \\
&= \int_0^T \int_\Omega \hat{h}\tilde{h}b_{noise} d\boldsymbol{x}\,dt - \int_0^T \int_\Omega \left(S\frac{\partial \tilde{q}}{\partial t} + \nabla\cdot e^{\gamma}\nabla\tilde{q} + \nabla\cdot\tilde{\gamma}e^{\gamma}\nabla q\right)\hat{h} d\boldsymbol{x}\,dt \\
&+ \int_\Omega S\tilde{q}(x,T)\hat{h}(x,T)d\boldsymbol{x} + \int_0^T \int_{\Gamma_N} e^{\gamma}\nabla\tilde{q}\cdot\boldsymbol{n}\hat{h} ds\,dt \\
&+ \int_0^T \int_{\Gamma_N} \tilde{\gamma}e^{\gamma}\nabla q\cdot\boldsymbol{n}\hat{h} ds\,dt
\end{aligned}$$

Thus, the incremental adjoint with boundary conditions is

$$\int_0^T \int_\Omega \hat{h}\tilde{h}b_{noise} d\boldsymbol{x}\,dt = \int_0^T \int_\Omega \left(S\frac{\partial \tilde{q}}{\partial t} + \nabla\cdot e^{\gamma}\nabla\tilde{q} + \nabla\cdot\tilde{\gamma}e^{\gamma}\nabla q\right)\hat{h} d\boldsymbol{x}\,dt$$

$$\int_\Omega S\tilde{q}(x,T)\hat{h}(x,T)d\boldsymbol{x} = 0$$

$$\int_0^T \int_{\Gamma_N} e^{\gamma}\nabla\tilde{q}\cdot\boldsymbol{n}\hat{h} ds\,dt = -\int_0^T \int_{\Gamma_N} \tilde{\gamma}e^{\gamma}\nabla q\cdot\boldsymbol{n}\hat{h} ds\,dt \Rightarrow e^{\gamma}\nabla\tilde{q}\cdot\boldsymbol{n} = 0 \text{ on } \Gamma_N$$

and the strong form of the expression is thus

$$\begin{aligned}
-S\frac{\partial \tilde{q}}{\partial t} - \nabla\cdot(e^{\gamma}\nabla\tilde{q}) &= \nabla\cdot(\tilde{\gamma}e^{\gamma}\nabla q) - \tilde{h}b_{\text{noise}} && \Omega\times(0,T) \\
\tilde{q}(x,T) &= 0 && \Omega\times\{t=T\} \\
\tilde{q} &= 0 && \Gamma_D\times(0,T) \\
e^{\gamma}\nabla\tilde{q}\cdot\boldsymbol{n} &= 0 && \Gamma_N\times(0,T)
\end{aligned}$$

3.2.2.3 Derivation of the reduced Hessian

Then the product of the reduced Hessian $\boldsymbol{H}$ with a vector $\tilde{\gamma} \in L^2(\Omega)$ is given by

$$\mathcal{L}_{\gamma h}(\tilde{h}) + \mathcal{L}_{\gamma q}(\tilde{q}) + \mathcal{L}_{\gamma\gamma}(\tilde{\gamma})$$

where $\tilde{h}$ and $\tilde{q}$ are the solutions to the incremental state and adjoint equations. This expression is built from the following variations:

$$\begin{aligned}
\mathcal{L}_{\gamma h}(\tilde{h}) &= \int_0^T \int_\Omega \hat{\gamma} e^\gamma \nabla \tilde{h} \cdot \nabla q d\boldsymbol{x} dt \\
\mathcal{L}_{\gamma q}(\tilde{q}) &= \int_0^T \int_\Omega \hat{\gamma} e^\gamma \nabla h \cdot \nabla \tilde{q} d\boldsymbol{x} dt \\
\mathcal{L}_{\gamma\gamma}(\tilde{\gamma}) &= \int_0^T \int_\Omega \hat{\gamma}\tilde{\gamma} e^\gamma \nabla h \cdot \nabla q d\boldsymbol{x} dt + R_{\gamma\gamma}
\end{aligned}$$

Taking into account boundary conditions, the discretized product of the reduced Hessian with a vector $\tilde{\gamma}$ is thus

$$\boldsymbol{C}_0(\boldsymbol{\gamma}, \tilde{\boldsymbol{h}}, \boldsymbol{q}) + \boldsymbol{C}_1(\boldsymbol{\gamma}, \tilde{\boldsymbol{q}}, \boldsymbol{h}) + (\boldsymbol{C}_2(\boldsymbol{\gamma}, \boldsymbol{q}, \boldsymbol{h}) + \boldsymbol{R})\tilde{\boldsymbol{\gamma}}$$

where

$$\begin{aligned}
\boldsymbol{C}_0(\gamma, \tilde{h}, q)_i &= \int_0^T \int_\Omega \psi_i e^\gamma \nabla \tilde{h} \cdot \nabla q d\boldsymbol{x} dt, \\
\boldsymbol{C}_1(\gamma, \tilde{q}, h)_i &= \int_0^T \int_\Omega \psi_i e^\gamma \nabla h \cdot \nabla \tilde{q} d\boldsymbol{x} dt, \\
\boldsymbol{C}_2(\gamma, q, h)_{i,j} &= \int_0^T \int_\Omega \psi_i \psi_j e^\gamma \nabla h \cdot \nabla q d\boldsymbol{x} dt.
\end{aligned}$$

3.3 Two parameter example

We study a model 1D problem in estimating the hydraulic conductivity based on observations of the head to provide a more concrete example of the algorithms described in Chapter 2. We first describe the relationship between the hydraulic conductivity and the head with a steady-state single phase flow model:

$$\begin{aligned} -\frac{d}{dx}\left(e^{\gamma}\frac{dh}{dx}\right) &= 80 \text{ in } (0,1) \\ h(0) &= 1 \\ h(1) &= 6, \end{aligned}$$

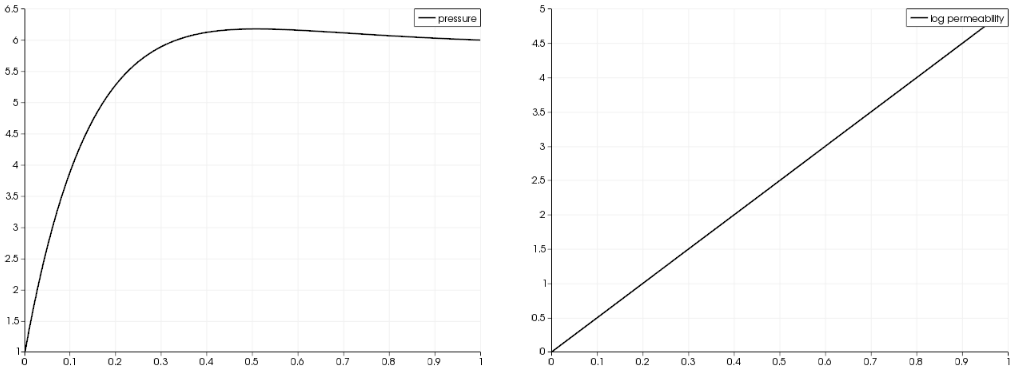
where h is the head and γ is the log hydraulic conductivity.

We discretize the problem with linear finite elements so that the hydraulic conductivity has 2 degrees of freedom (so that we can easily visualize the results of this model problem) while the head has 129 degrees of freedom. We generate synthetic observations from a problem with an analytical solution where the head is given by the expression

$$\begin{aligned} h(x) &= \frac{e^{g_0}g_0h_1 - g_0h_0 - c_{rhs} + e^{-g_0x}(c_{rhs}(1+(e^{g_0}-1)x) + e^{g_0}g_0(h_0-h_1))}{g_0(-1+e^{g_0})} \\ &= \frac{-17+6e^5+e^{-5x}(16-5e^5+16(e^5-1)x)}{-1+e^5} \\ &\approx 5.92538 + e^{-5x}(-4.92538+16x) \end{aligned} \tag{3.4}$$

with parameters $g_0 = 5$, $h_0 = 1$, $h_1 = 6$, $c_{rhs} = 80$, $g_1 = 0$, and $\gamma = g_0x$, and can be seen in Figure 3.1.

We see in Figure 3.2 that the likelihood is nonlinear and depends on the location of the observation points (as illustrated by Figure 3.2(a)-3.2(c)). Thus, even though the prior pdf (in Figure 3.2(d)) is Gaussian, the posterior pdf is non-Gaussian.



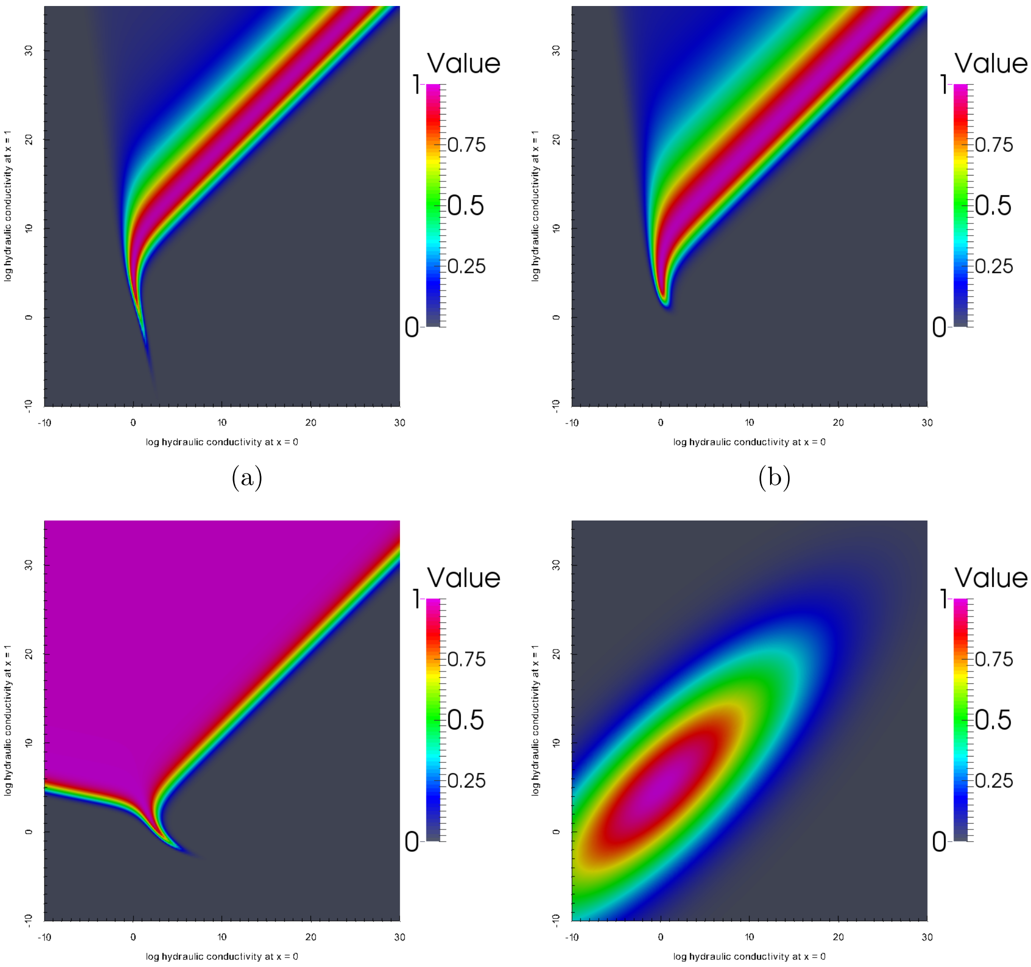
(a) (b)

Figure 3.1: Synthetic data for example problem given in Equation 3.4: (a) head, with four observations $(x, y_{obs}) = (0.1, 3.908)$, $(0.2, 5.290)$, $(0.8, 6.070)$, $(0.9, 6.031)$. (b) log hydraulic conductivity $\gamma = 5x$.

3.3.1 Approximation of $\pi_{\mathbf{post}}(\boldsymbol{x})$

In this section, we visualize the steps of the algorithm in §2.2 to build the Hessian-based response surfaces. Since there are only two parameters, a low-rank approximation to the Hessian is unnecessary. We do, however, need a positive definite approximation to the Hessian of the data misfit for the purpose of building the Hessian-based response surfaces. Here we present results using the Gauss-Newton approximation to the Hessian. In later chapters we simply restrict our approximation to positive eigenvalues of the prior-preconditioned Hessian of the data misfit (this is the approach used in Chapter 6).

Our goal is approximation of the (non-Gaussian) posterior pdf in Figure 3.3(a) (which, by Bayes' theorem, is equal to Figure 3.2(b) times Figure 3.2(d)). For comparison, in Figure 3.3 we also show the Gaussian process response surface (restricted to non-negative values) and the Gaussian mixture



(a) (b)

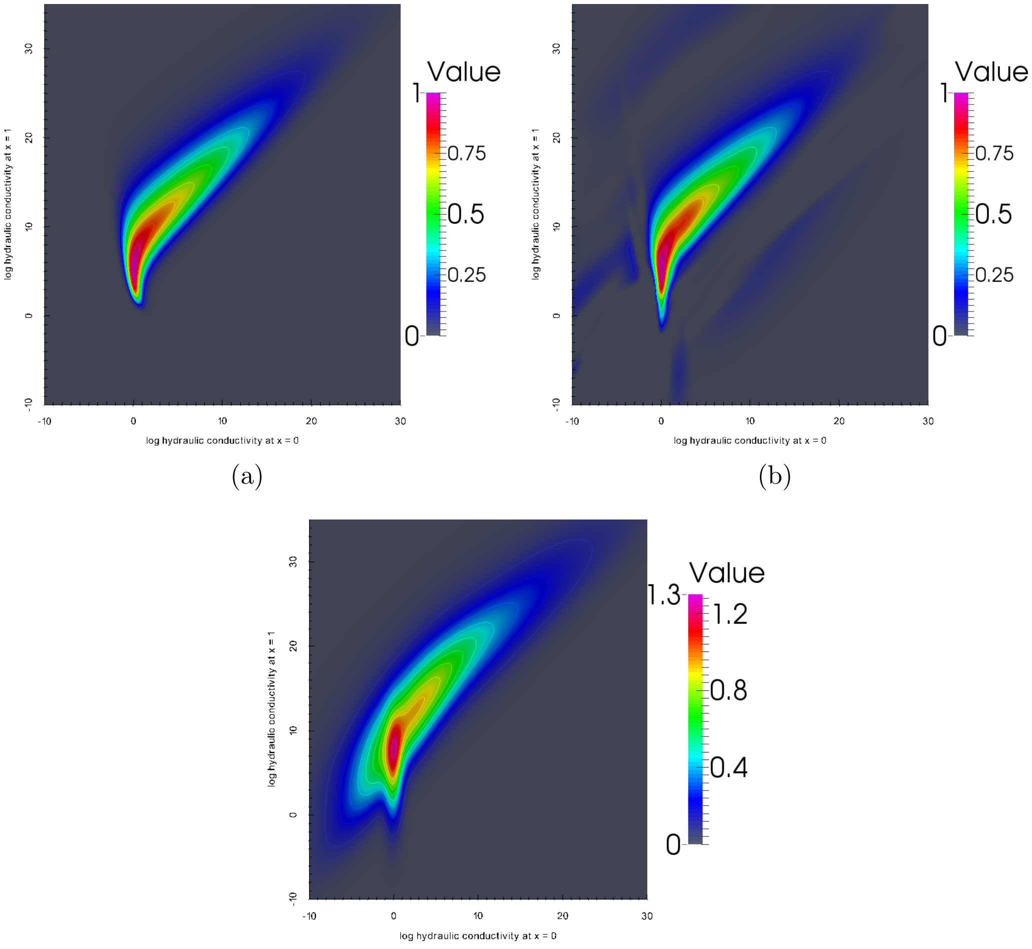
(c) (d)

Figure 3.2: Likelihood for sensor observations at (a) x = 0.1, 0.15, 0.2, 0.25 (b) x = 0.1, 0.2, 0.8, 0.9 (c) x = 0.75, 0.8, 0.85, 0.9. (d) Prior pdf with covariance parameters $\beta_{0,\text{prior}} = 0.1$ and $\beta_{1,\text{prior}} = 0.005$

obtained after 30 greedy iterations. In the MCMC algorithms proposed in §2.3, the Gaussian mixture is employed as a proposal, and the GPRS serves as a surrogate for the posterior pdf. We see that the GPRS (in Figure 3.3(b)) is a better approximation to the posterior pdf than the Gaussian mixture (in Figure 3.3(c)), as expected since it is the result of an interpolation between a set of training points. However, it is easy to draw samples from the Gaussian mixture so it is a good candidate for a proposal since it still respects the shape of the posterior pdf.

We present the evolution of the GPRS and Gaussian mixture step by step in Figures 3.4-3.6. In the first step of the greedy iteration, we find the MAP point. The Gaussian mixture (shown in Figure 3.4(a)) is simply a Gaussian at the MAP point with a covariance equal to the inverse of the Hessian. Note that using this approximation would be highly misleading. The GPRS shown in Figure 3.4(b) is an interpolation including the Hessian at the MAP point and additional interpolation information at the value points along the path of the optimization algorithm. The additional information leads to some negative oscillations in other regions, but the response surface conforms to the values at these points.

After the second greedy iteration, we add a new shape point. The Gaussian mixture at this iteration is shown in Figure 3.4(c). Compared to the exact posterior, the values are too high to the left of the MAP point and the tail does not extend far enough to the right. This is partially corrected in the Gaussian process response surface in Figure 3.4(d) by the additional



(a)

(b)

(c)

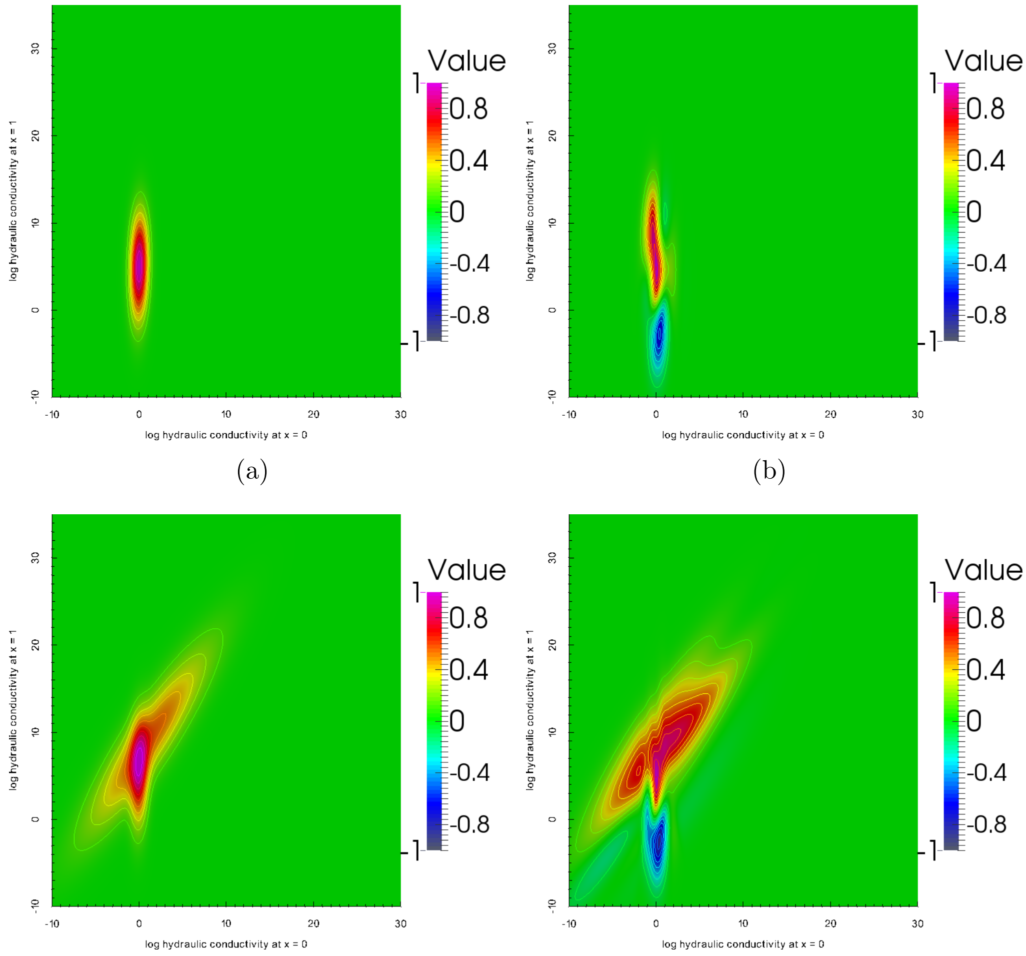
Figure 3.3: (a) Posterior pdf. (b) non-negative Gaussian process response surface after 30 greedy iterations. (c) Gaussian mixture after 30 greedy iterations.

information at the value points.

The third greedy iteration (in Figure 3.5(a)-3.5(b)) finds a small local error maximum where the exact pdf is near zero but the response surface is not. It simply adds value points to correct the interpolation, not a shape point. Since there is no new shape point, the Gaussian mixture in Figure 3.5(a) is unchanged. The Gaussian process response surface only changes slightly in Figure 3.5(b). The fourth greedy iteration also only adds value points (no shape point), but the effect on the response surface in Figure 3.5(d) is more dramatic. The new value points are to the left of the MAP point, eliminating the overshoot seen in Figure 3.5(b), but the response surface then bulges to the right.

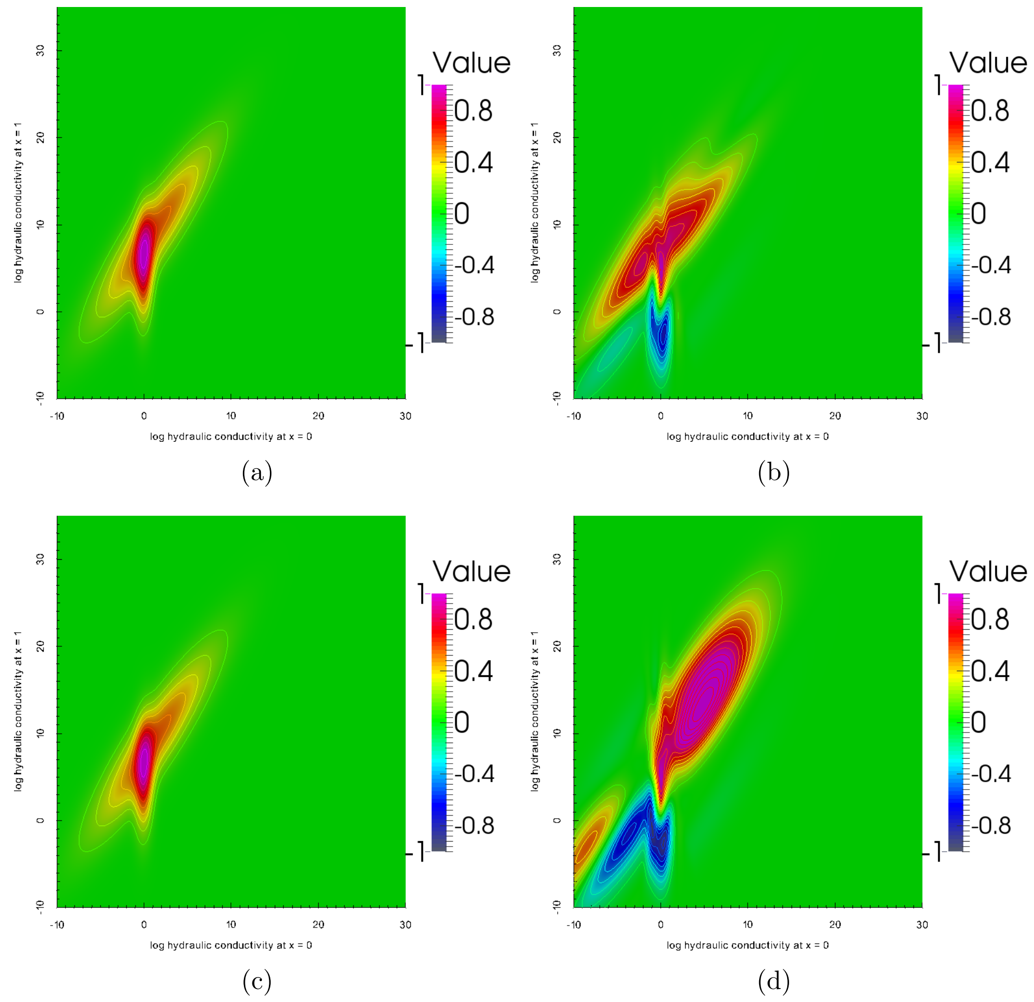
Continuing in Figure 3.6 with the fifth greedy iteration, we again only add value points. Comparing Figure 3.5(d) to Figure 3.6(b), it is apparent that these remove the positive oscillation to the bottom left of the MAP point. In greedy iteration 6, we add a new shape point. The Gaussian mixture in Figure 3.6(c) extends further to the right than in the previous iteration, and the bulge in the response surface has been reduced in Figure 3.6(d).

We now skip ahead to examine the Gaussian process response surface after 10, 20, and 30 greedy iterations, seen in Figure 3.7. The changes are less dramatic than the first 6 iterations, but they are still visible. Note that at 10 iterations, the approximation still has too short of a tail to the right and several large negative oscillations, but is still a far better approximation than a Gaussian at the MAP point.



(a) (b) (c) (d)

Figure 3.4: Colormap on a scale from -1 to 1, with red corresponding to positive values and blue corresponding to negative values. (a) Gaussian at the MAP point (after 1 greedy iteration). (b) GPRS after 1 greedy iteration. (c) Gaussian mixture after 2 greedy iterations. (d) GPRS after 2 greedy iterations.



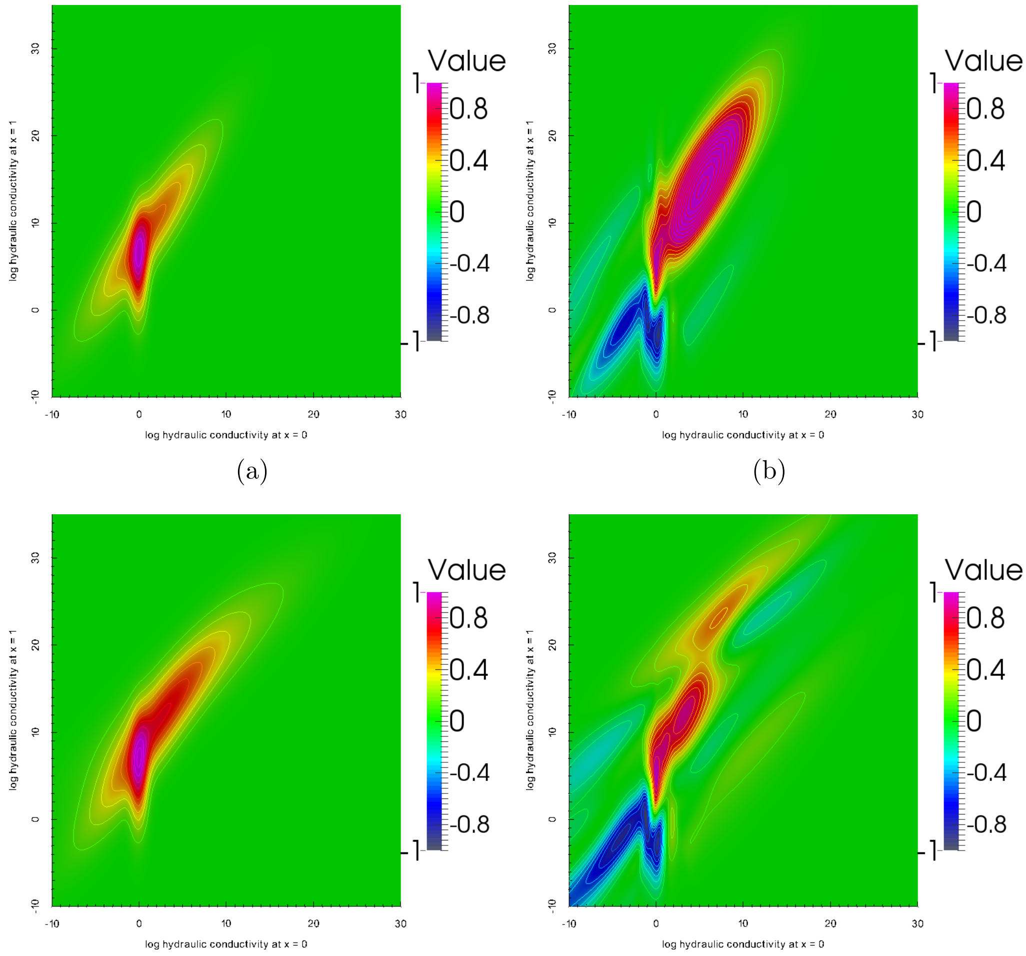
(a)

(b)

(c)

(d)

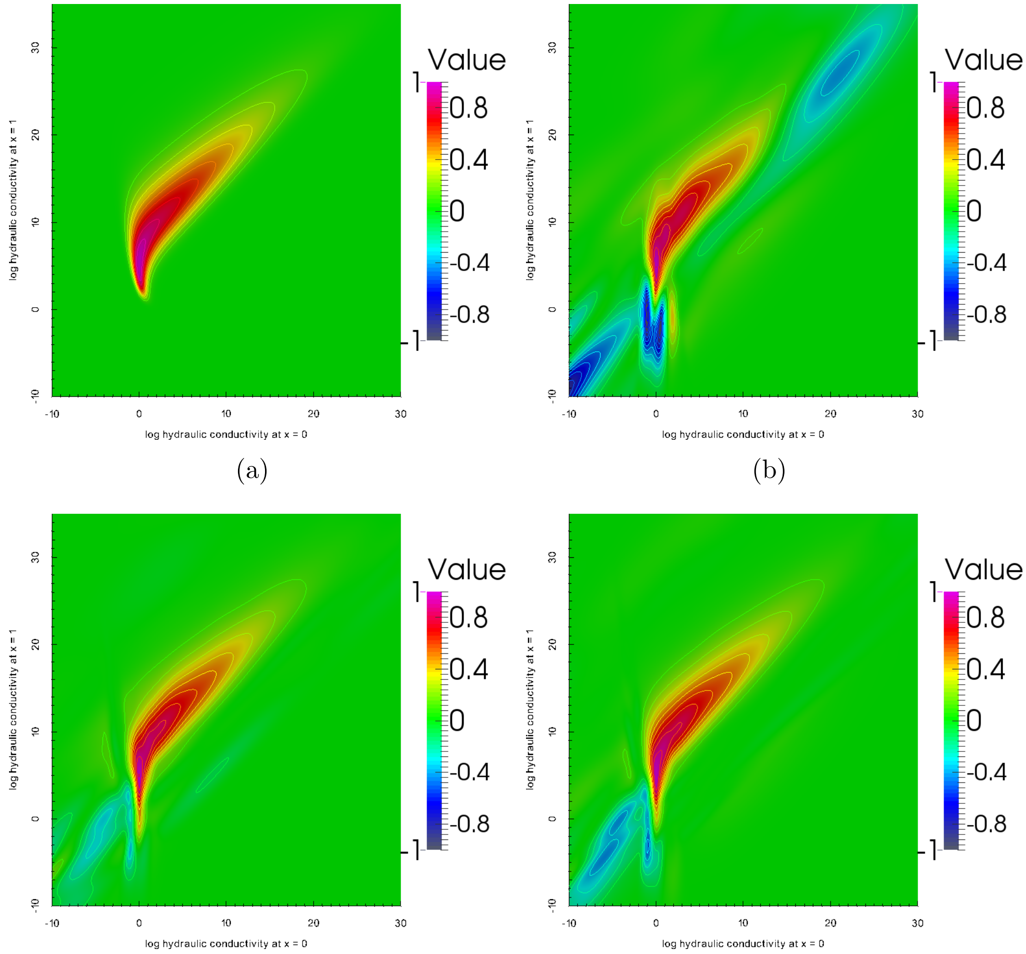
Figure 3.5: Colormap on a scale from -1 to 1, with red corresponding to positive values and blue corresponding to negative values. (a) Gaussian mixture after 3 greedy iterations. (b) GPRS after 3 greedy iterations. (c) Gaussian mixture after 4 greedy iterations. (d) GPRS after 4 greedy iterations.



(a) (b)

(c) (d)

Figure 3.6: Colormap on a scale from -1 to 1, with red corresponding to positive values and blue corresponding to negative values. (a) Gaussian mixture after 5 greedy iterations. (b) GPRS after 5 greedy iterations. (c) Gaussian mixture after 6 greedy iterations. (d) GPRS after 6 greedy iterations.



(a) (b)

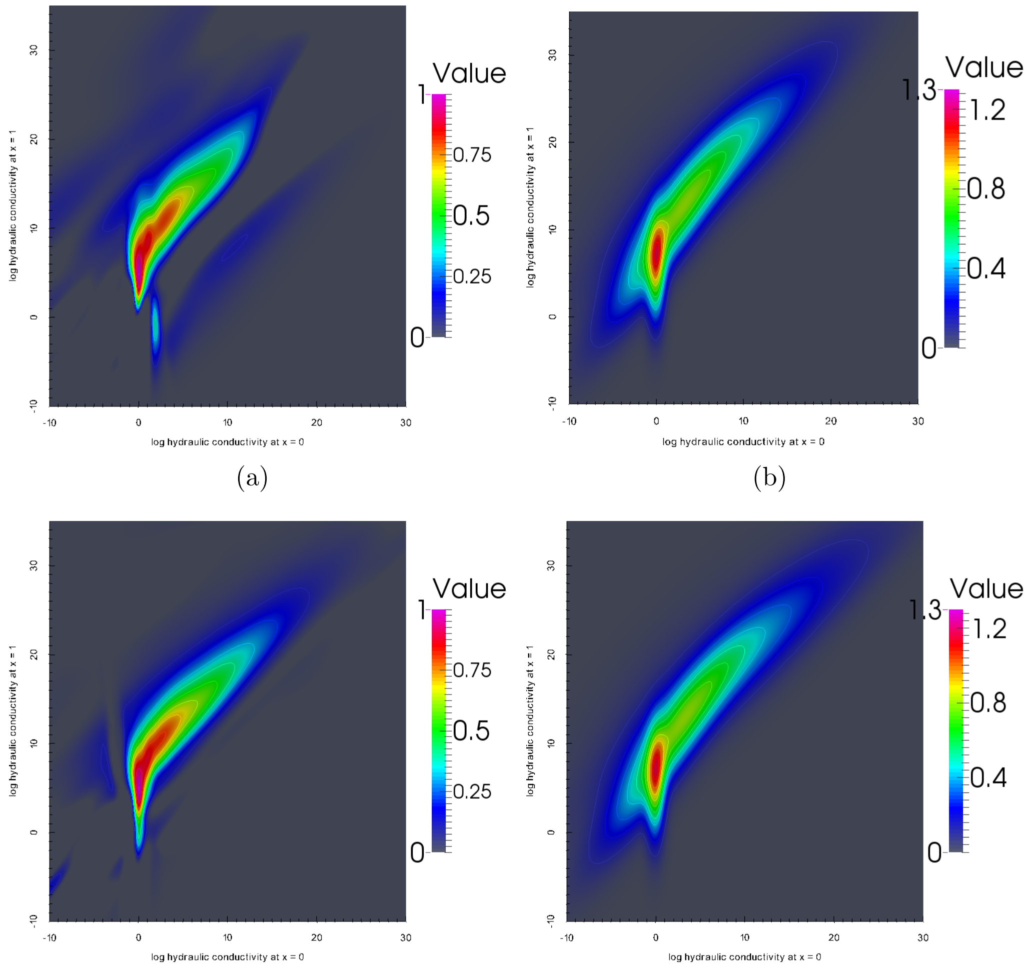
(c) (d)

Figure 3.7: (a) Posterior pdf. (b) Gaussian process response surface after 10 greedy iterations. (c) Gaussian process response surface after 20 greedy iterations. (d) Gaussian process response surface after 30 greedy iterations.

To study the effect of the number of greedy iterations on the GPRS, we examine the Gaussian mixture and the GPRS restricted to non-negative values after 10 and 20 greedy iterations in Figure 3.8. If the greedy algorithm is terminated too early then the GPRS may not be an accurate approximation for the posterior pdf – the tail is still too short after 10 greedy iterations. We also see that the Gaussian mixture respects the general structure of the posterior pdf.

In Figure 3.9 we show the development of the chain generated by Algorithm 1 in §2.3 (Metropolis-Hastings sampling of the posterior pdf with the Gaussian mixture as a proposal). We show that with additional greedy iterations, the Gaussian mixture expands from a Gaussian at the MAP point (Figure 3.9(a)) to include the tail of the posterior pdf (Figures 3.9(b)-3.9(d)).

To provide a visual comparison of the chains produced by all three MCMC algorithms (from §2.3), in Figure 3.10 and Figure 3.11 we show the MCMC chains generated using the proposed algorithms after 6 greedy iterations and 30 greedy iterations respectively. The differences are particularly visible in Figure 3.10, where the GPRS (Figure 3.10(a)) is still a poor approximation to the posterior pdf. Therefore the differences in the chain shown in Figure 3.10(c), where the GPRS serves as a surrogate for the posterior pdf, and the chain in Figure 3.10(c) (the two-stage delayed acceptance algorithm) are particularly noticeable. However, as we see in Figure 3.11, the visible differences between the algorithms decrease as additional greedy iterations improve the response surfaces.

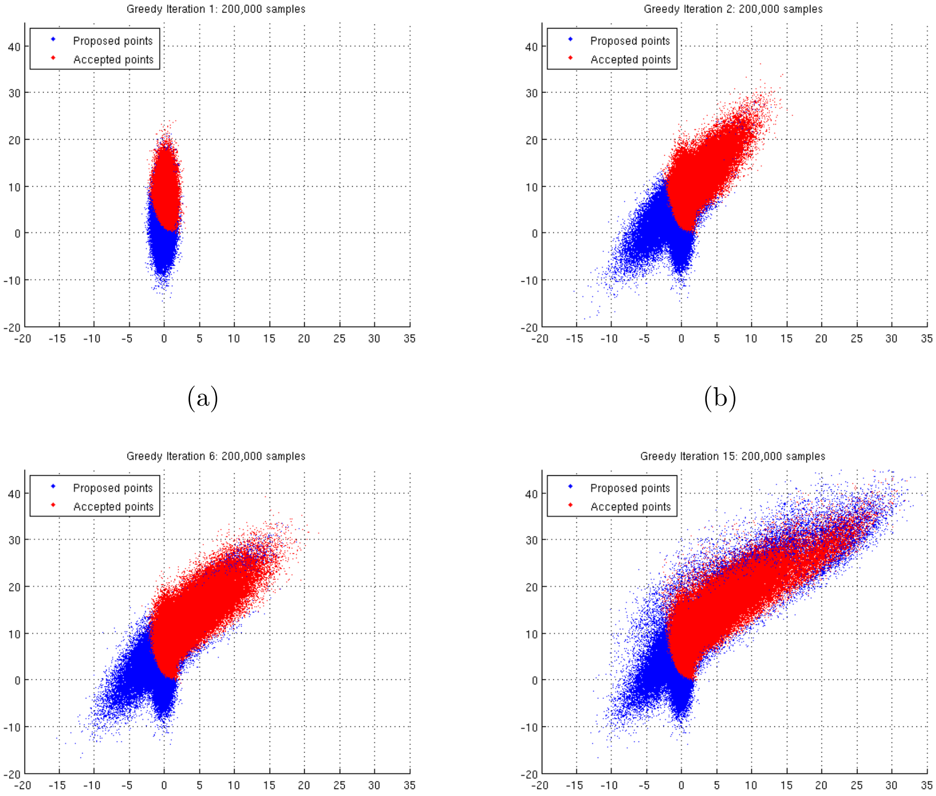


(a) (b)

(c) (d)

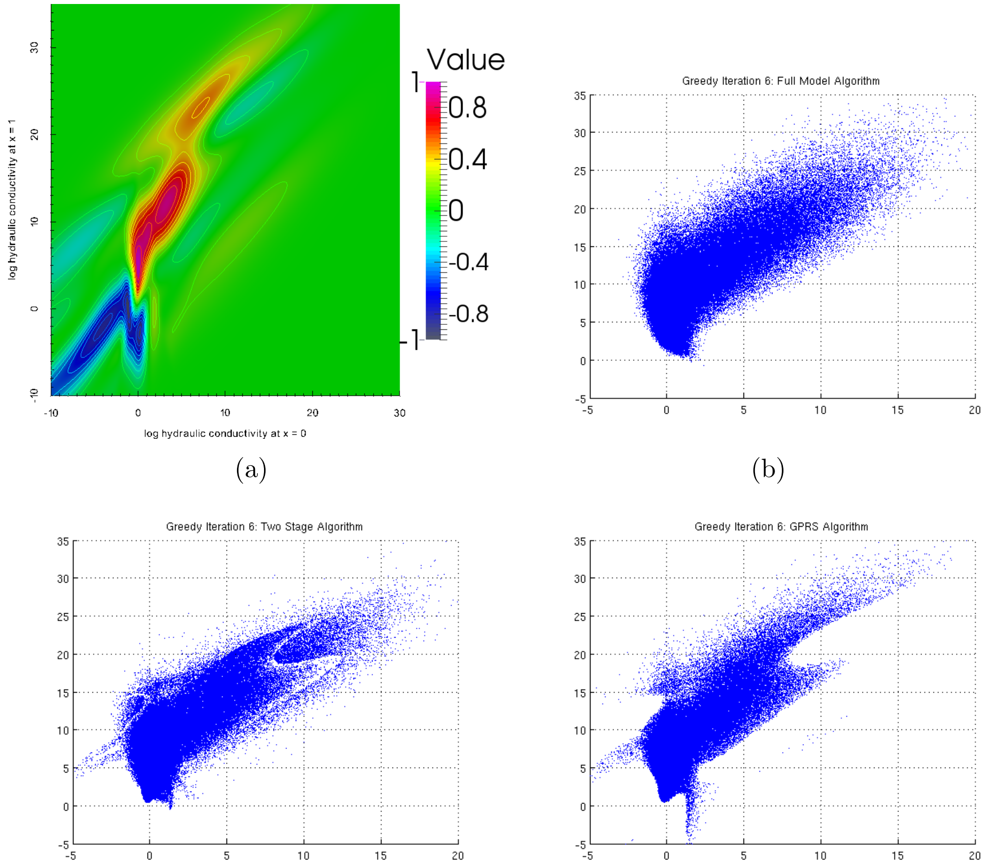
Figure 3.8: (a) GPRS restricted to non-negative values, computed through 10 greedy iterations. (b) Gaussian mixture after 10 greedy iterations. (c) GPRS after 20 greedy iterations, restricted to non-negative values. (d) Gaussian mixture after 20 greedy iterations

In conclusion, the GPRS is a good candidate to act as a surrogate for the posterior pdf, and the Gaussian mixture is a good candidate for a proposal. Full numerical results for a 121 parameter problem are in Chapter 6.



(a) (b) (c) (d)

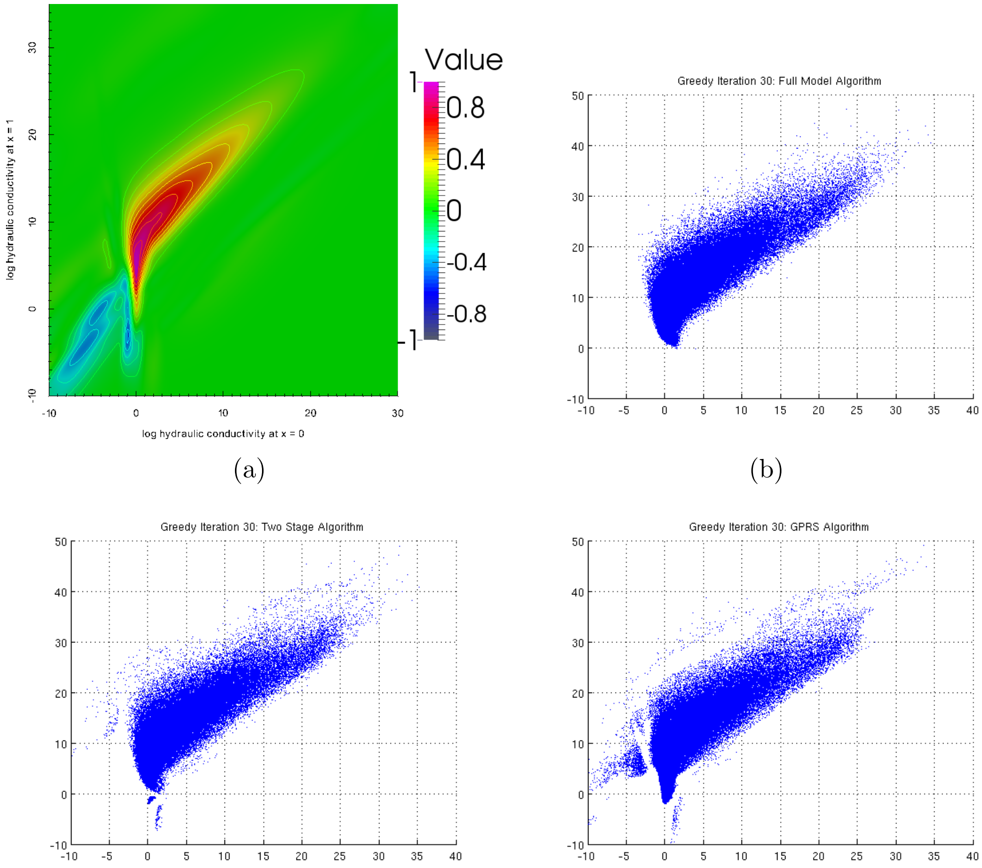
Figure 3.9: In blue, samples drawn from the Gaussian mixture. In red, points accepted during Metropolis-Hastings sampling of the posterior pdf with the Gaussian mixture as a proposal after (a) 1 greedy iteration, (b) 2 greedy iterations, (c) 6 greedy iterations, (d) 15 greedy iterations.



(a) (b)

(c) (d)

Figure 3.10: (a) The Gaussian process response surface after 6 greedy iterations. (b) Chain from Metropolis-Hastings sampling of the posterior pdf with the Gaussian mixture proposal after 6 greedy iterations. (c) Chain from the two-stage delayed acceptance algorithm after 6 greedy iterations. (d) Chain from Metropolis-Hastings sampling of the GPRS as a surrogate for the posterior pdf, with the Gaussian mixture proposal, after 6 greedy iterations.



(a) (b) (c) (d)

Figure 3.11: (a) The Gaussian process response surface after 30 greedy iterations. (b) Chain from Metropolis-Hastings sampling of the posterior pdf with the Gaussian mixture proposal after 6 greedy iterations. (c) Chain from the two-stage delayed acceptance algorithm after 30 greedy iterations. (d) Chain from Metropolis-Hastings sampling of the GPRS as a surrogate for the posterior pdf, with the Gaussian mixture proposal, after 30 greedy iterations.

Chapter 4

Low-rank approximation of the Hessian

We next address the scalability of our response surface approximation (and our Gaussian mixture approximation) to high-dimensional parameter spaces. The expressions for $\mu_{\mathrm{GPRS}}(\boldsymbol{x})$ and $\Gamma_{\mathrm{GPRS}}(\boldsymbol{x})$ rely on evaluating the product of the Hessian at each training point in $\boldsymbol{M}_k$ with a vector determined by the value of $\boldsymbol{x}$. However, computing a Hessian-vector product is not cheap. Since the goal is to produce an approximation that can be sampled many times in order to explore a high-dimensional parameter space, the scalability of the algorithm requires exploiting the structure of the Hessian so that the Hessian may be replaced by a low-rank approximation. A faithful low-rank approximation may often be computed at a cost independent of the size of the parameter space. For large-scale problems, the cost of computing the low-rank approximation is much less than that of evaluating Hessian-vector products required to probe the parameter space. Thus, there is a scalable way to incorporate the structure of the underlying statistical inverse problem into our response surface approximation of the posterior pdf for high-dimensional parameter spaces.

The Hessian

$$\boldsymbol{H} = -\nabla^2(\log \pi_{\text{post}}(\boldsymbol{x})), \tag{4.1}$$

where $\pi_{\text{post}}(\boldsymbol{x})$ is described by (2.1), is fundamental to the construction of our response surface approximation of the posterior pdf. This ties the response surface approximation to the underlying statistical inverse problem and to the classical deterministic least-squares optimization formulation.

Although it is easy to write an expression such as (4.1), explicit construction of the Hessian, which is dense and of order of the number of parameters, would require at least as many forward solves as the number of parameters. This is prohibitive for high-dimensional parameter spaces, as may arise when an infinite-dimensional parameter field is discretized, and expensive-to-solve PDEs. Instead, we take advantage of the structure of the Hessian. Combining (4.1) with (2.1), we note that the Hessian is composed of two terms: the Hessian of the data misfit $\boldsymbol{H}_{\text{misfit}} \in \mathbb{R}^{n\times n}$, where

$$\boldsymbol{H}_{\text{misfit}} \overset{\text{def}}{=} \nabla^2 \left(\frac{1}{2} \|\boldsymbol{y}_{\text{obs}} - f(\boldsymbol{x}) - \bar{\boldsymbol{e}}\|_{\boldsymbol{\Gamma}^{-1}_{\text{noise}}} \right),$$

and the Hessian of the term resulting from the Gaussian prior pdf. The Hessian can therefore be expressed as

$$\boldsymbol{H} = \boldsymbol{H}_{\text{misfit}} + \boldsymbol{\Gamma}^{-1}_{\text{prior}}. \tag{4.2}$$

For many ill-posed inverse problems, the spectrum of $\boldsymbol{H}_{\text{misfit}}$ (the data misfit portion of the Hessian) decays rapidly, reflecting the fact that the data provide information on only a low-dimensional subspace of the parameter field. One would like to employ a low-rank approximation of this operator (e.g., via a Lanczos method) to overcome the prohibitive nature of direct computation. However, the retained modes should be informed by not only the data, but also the prior information, since modes that can be inferred from the data might be nullified by the influence of the prior. Therefore, we rearrange the expression to include the effects of the prior in the data misfit term, thereby filtering the data through the prior before invoking the low-rank approximation. We argue in subsequent sections that for many ill-posed inverse problems, an accurate low-rank spectral approximation of the "prior-preconditioned Hessian of the data misfit," $\mathbf{\Gamma}_{\text{prior}}^{1/2}\boldsymbol{H}_{\text{misfit}}\mathbf{\Gamma}_{\text{prior}}^{1/2}$ can be computed at a cost that is a small multiple (independent of the parameter dimension) of the cost of solving the underlying forward PDEs.

Much of the prior work for low-rank approximations in inverse problems has occurred in the context of seismic inverse problems, where the forward operators can be expensive, and the heterogeneity of the earth requires a high-dimensional parametrization. When the forward problem is sufficiently inexpensive to solve, one can explicitly construct the Hessian matrix (at a cost of n forward problem solutions, where n is the number of model parameters) and (for a linear problem) invert it to obtain the posterior covariance matrix. For example, as early as 1993, [54] were able to explicitly compute the posterior covariance in the Bayesian framework for a global seismic tomography problem using ray tracing as the forward model for up to 12,496 model parameters. Similarly, when the parametrization is low-dimensional, one can tractably construct and invert the Hessian matrix to find the posterior covariance. For example, [28] and [29] obtain Bayesian solutions to flat-layered earth seismic inverse problems with up to 300 parameters. Such explicit construction of the Hessian does not scale as the number of parameters, or the complexity of the forward problem, increase. On the other hand, a number of references invoke low-rank approximations to make inversion tractable for large-scale problems. Examples include [46, 27, 11, 12, 42, 53, 59, 57, 47, 13, 58], in which the low rank approximations are effected through truncated iterative methods such as Lanczos and block Lanczos, conjugate gradients, and LSQR. In all of these cases, however, the Bayesian framework is not employed, and instead the goal is to determine the so-called resolution matrix $\boldsymbol{A}^{\dagger}\boldsymbol{A}$ [3], which is a deterministic attempt to quantify resolving power of the inversion.

4.1 Theory: Approximation of the Hessian matrix

We begin by recognizing that, for many ill-posed inverse problems, the Hessian of the data misfit, $\boldsymbol{H}_{\text{misfit}}$, behaves like the discretization of a compact operator. The range space thus is effectively finite-dimensional, and the eigenvalues decay, often rapidly, to zero. We can exploit this structure to construct fast algorithms for approximating the Hessian. Rearranging the expression for $\boldsymbol{H}$ in (4.2) to factor out $\boldsymbol{\Gamma}_{\text{prior}}^{1/2}$ gives

$$\boldsymbol{H} = \boldsymbol{\Gamma}_{\text{prior}}^{-1/2} \left(\boldsymbol{\Gamma}_{\text{prior}}^{1/2} \boldsymbol{H}_{\text{misfit}} \boldsymbol{\Gamma}_{\text{prior}}^{1/2} + \boldsymbol{I} \right) \boldsymbol{\Gamma}_{\text{prior}}^{-1/2}. \tag{4.3}$$

Using this Hessian within the Gaussian process approximation requires two operations: fast multiplication by $\boldsymbol{\Gamma}_{\text{prior}}^{-1/2}$ and fast multiplication by the matrix

$\boldsymbol{\Gamma}_{\text{prior}}^{1/2}\boldsymbol{H}_{\text{misfit}}\boldsymbol{\Gamma}_{\text{prior}}^{1/2} + \boldsymbol{I}$. We do not address fast multiplication by $\boldsymbol{\Gamma}_{\text{prior}}^{1/2}$ here, since it depends on the structure of the chosen prior. In any case, it is independent of the PDE solve, and is thus relatively cheap. Here we address the latter operation.

Let λ_i and $\boldsymbol{v}_i$ be the eigenvalues and eigenvectors of the *prior-preconditioned Hessian of the data misfit* $\tilde{\boldsymbol{H}}_{\text{misfit}} \in \mathbb{R}^{n\times n}$, where

$$\tilde{\boldsymbol{H}}_{\text{misfit}} \overset{\text{def}}{=} \boldsymbol{\Gamma}_{\text{prior}}^{1/2}\boldsymbol{H}_{\text{misfit}}\boldsymbol{\Gamma}_{\text{prior}}^{1/2}.$$

Let $\boldsymbol{\Lambda} \in \mathbb{R}^{n\times n}$ be the diagonal matrix of eigenvalues λ_i of $\tilde{\boldsymbol{H}}_{\text{misfit}}$ and let $\boldsymbol{V} \in \mathbb{R}^{n\times n}$ be the matrix whose columns are the eigenvectors $\boldsymbol{v}_i$ of $\tilde{\boldsymbol{H}}_{\text{misfit}}$. Then the following expression can be rewritten with $\tilde{\boldsymbol{H}}_{\text{misfit}}$ replaced by its spectral decomposition:

$$\boldsymbol{\Gamma}_{\text{prior}}^{1/2}\boldsymbol{H}_{\text{misfit}}\boldsymbol{\Gamma}_{\text{prior}}^{1/2} + \boldsymbol{I} = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^T + \boldsymbol{I}$$

When the eigenvalues of $\tilde{\boldsymbol{H}}_{\text{misfit}}$ decay rapidly, we can extract a low-rank approximation of $\tilde{\boldsymbol{H}}_{\text{misfit}}$ by retaining only the r largest eigenvalues and corresponding eigenvectors,

$$\boldsymbol{\Gamma}_{\text{prior}}^{1/2}\boldsymbol{H}_{\text{misfit}}\boldsymbol{\Gamma}_{\text{prior}}^{1/2} \approx \boldsymbol{V}_r\boldsymbol{\Lambda}_r\boldsymbol{V}_r^T.$$

and therefore

$$\boldsymbol{\Gamma}_{\text{prior}}^{1/2}\boldsymbol{H}_{\text{misfit}}\boldsymbol{\Gamma}_{\text{prior}}^{1/2} + \boldsymbol{I} = \boldsymbol{V}_r\boldsymbol{D}_r\boldsymbol{V}_r^T + \boldsymbol{I} + \mathcal{O}\left(\sum_{i=r+1}^{n} \lambda_i^2\right), \tag{4.4}$$

where $\boldsymbol{\Lambda}_r \in \mathbb{R}^{r\times r}$ and $\boldsymbol{V}_r \in \mathbb{R}^{n\times r}$ denote the truncated eigenvalue and eigenvector matrices. With this low-rank approximation, the final expression for approximation of the Hessian is given by

$$\boldsymbol{H} \approx \boldsymbol{\Gamma}_{\text{prior}}^{-1/2} \boldsymbol{V}_r \boldsymbol{\Lambda}_r \boldsymbol{V}_r^T \boldsymbol{\Gamma}_{\text{prior}}^{-1/2} + \boldsymbol{\Gamma}_{\text{prior}}^{-1}. \tag{4.5}$$

For many ill-posed inverse problems, the choice of r is small and independent of problem size. The r retained eigenvectors can be viewed as the modes of the parameter field that are recoverable from a combination of data and prior information. It is often the case for ill-posed inverse problems that the spectrum of the prior-preconditioned Hessian of the data misfit $\tilde{\boldsymbol{H}}_{\text{misfit}}$ collapses on zero. This happens for example when the data misfit Hessian $\boldsymbol{H}_{\text{misfit}}$ has rapidly decaying eigenvalues, and/or when the chosen prior is of smoothing type. In such cases, an accurate approximation of $\tilde{\boldsymbol{H}}_{\text{misfit}}$ can be made with small values of r. In fact, often the data and prior are informative about the low-wavenumber modes of the parameter field; this is because local features cannot be resolved from the data, and the prior is of smoothing type. Thus, refinements of the mesh on which the PDE is solved do not affect the accuracy of the low-rank approximation, and as a result, an appropriate choice of r is independent of mesh size. Note that expression (4.4) suggests a cutoff value for retaining eigenpairs in the low-rank approximation of $\tilde{\boldsymbol{H}}_{\text{misfit}}$, namely that $\lambda \ll 1$ (in most of our examples, we use $\lambda > 0.1$ as a cutoff criterion).

The Frobenius norm of the error in the approximation of $\tilde{\boldsymbol{H}}_{\text{misfit}}$ is

$$\left\| \tilde{\boldsymbol{H}}_{\text{misfit}} - \boldsymbol{V}_r \boldsymbol{\Lambda}_r \boldsymbol{V}_r^T \right\|_F = \sqrt{\sum_{j=r+1}^{n} \lambda_j^2}. \tag{4.6}$$

If the eigenvalues λ_j decay rapidly enough for the series $\sum_{j=r+1}^{\infty} \lambda_j^2$ to converge, then provided the structure of the spectrum does not change as the dimension of the parameter space increases, we obtain an error bound on the low-rank approximation of $\tilde{\boldsymbol{H}}_{\text{misfit}}$ that is independent of the dimension of the parameter space (and therefore mesh size). Finally, to relate the effect of the low-rank approximation of $\tilde{\boldsymbol{H}}_{\text{misfit}}$ on the full Hessian $\boldsymbol{H}$, one would need to know the spectral structure of the prior covariance $\boldsymbol{\Gamma}_{\text{prior}}^{-1}$, as is evident from (4.3).

We are now in a position to state the complexity of the low-rank approximation and the resulting cost of evaluating $\boldsymbol{H}$. We certainly cannot explicitly construct $\tilde{\boldsymbol{H}}_{\text{misfit}}$ and then perform a truncated SVD. Instead, we chose a matrix-free Lanczos method to find its dominant eigenvalues and corresponding eigenvectors. Lanczos requires only a matrix-vector product with $\tilde{\boldsymbol{H}}_{\text{misfit}}$ at each iteration, which in turn requires (a) multiplications with $\boldsymbol{\Gamma}_{\text{prior}}^{1/2}$, and (b) action of $\boldsymbol{H}_{\text{misfit}}$ on vectors. The latter involves forward and adjoint PDE solutions, which dominate the cost when the forward model is large-scale. Lanczos tends to require a number of iterations proportional to the dominant portion of the spectrum, which is bounded when $\tilde{\boldsymbol{H}}_{\text{misfit}}$ is compact. Thus, under the conditions discussed above, the cost (in number of forward/adjoint PDE model solutions) of the low-rank approximation is small and independent of the mesh size. Once the r dominant eigenpairs have been found, storing the low-rank approximation requires $(r+1)n$ floating point numbers, where n is the dimension of the parameter space. Once the low-rank approximation has been constructed, the cost of a Hessian-vector multiplication is $4rn$ plus the cost of two matrix-vector multiplications with $\mathbf{\Gamma}_{\text{prior}}^{-1/2}$.

In §5.2 and §5.3 we illustrate the relationship between accuracy and size of r and provide evidence of mesh independence for a large-scale statistical inverse problem governed by a three-dimensional convection-diffusion PDE model. In §4.3.1 and §4.3.2 we provide similar results for a groundwater example.

4.2 Local approximation of an inverse problem

We address the question of linearizing the inverse problem around a point because the Gaussian mixture in §2.2.4 is a sum of local linear approximations around the shape points $\boldsymbol{x}_k$.

If the posterior pdf $\pi_{\text{post}}(\boldsymbol{x})$ is approximated by a linearization at a shape point $\boldsymbol{x}_k$ (chosen through the greedy algorithm described in Chapter 2), then locally

$$\pi_{\text{post}}(\boldsymbol{x}) \approx \mathcal{N}(\bar{\boldsymbol{x}}_k, \mathbf{\Gamma}_k),$$

and this posterior covariance matrix is the inverse of the Hessian at the point $\boldsymbol{x}_k$,

$$\mathbf{\Gamma}_k = \left(\boldsymbol{H}_{\text{misfit}} + \mathbf{\Gamma}_{\text{prior}}^{-1}\right)^{-1}. \tag{4.7}$$

In the linear case (such as in Chapter 5), this approximation is exact since $\pi_{\text{post}}(\boldsymbol{x})$ is also Gaussian, with mean $\bar{\boldsymbol{x}}_{\text{post}} \in \mathbb{R}^n$ given by the maximum *a posteriori* (MAP) point, i.e.,

$$\bar{\boldsymbol{x}}_{\text{post}} = \arg\max_{\boldsymbol{x}} \pi_{\text{post}}(\boldsymbol{x}).$$

4.2.1 Low rank approximation of the posterior covariance

We now turn to the problem of approximating the posterior covariance matrix of the parameters and its square root with determinant, which are necessary for drawing a sample from the Gaussian mixture used in all the Hessian-based MCMC algorithms described in §2.3.

As discussed already, explicit construction of the Hessian (let alone its inverse), is prohibitive for high-dimensional parameter spaces. Moreover, for large-scale problems we often cannot store the entire posterior covariance matrix, and instead must resort to extracting partial information (such as its action on a vector or the variance). Instead, we will take advantage of the low-rank approximation of $\tilde{\boldsymbol{H}}_{\text{misfit}}$ described in §4.1. Rearranging the expression for $\boldsymbol{\Gamma}_{\text{post}}$ in (4.7) to factor out $\boldsymbol{\Gamma}_{\text{prior}}^{1/2}$ gives

$$\boldsymbol{\Gamma}_{\text{post}} = \boldsymbol{\Gamma}_{\text{prior}}^{1/2} \left(\boldsymbol{\Gamma}_{\text{prior}}^{1/2} \boldsymbol{H}_{\text{misfit}} \boldsymbol{\Gamma}_{\text{prior}}^{1/2} + \boldsymbol{I} \right)^{-1} \boldsymbol{\Gamma}_{\text{prior}}^{1/2}. \tag{4.8}$$

As in §4.1, we extract a low-rank approximation of $\tilde{\boldsymbol{H}}_{\text{misfit}}$ by retaining only the r largest eigenvalues and corresponding eigenvectors,

$$\boldsymbol{\Gamma}_{\text{prior}}^{1/2} \boldsymbol{H}_{\text{misfit}} \boldsymbol{\Gamma}_{\text{prior}}^{1/2} \approx \boldsymbol{V}_r \boldsymbol{\Lambda}_r \boldsymbol{V}_r^T.$$

Then we can invert using the Sherman-Morrison-Woodbury formula to obtain

$$\left(\boldsymbol{\Gamma}_{\text{prior}}^{1/2} \boldsymbol{H}_{\text{misfit}} \boldsymbol{\Gamma}_{\text{prior}}^{1/2} + \boldsymbol{I} \right)^{-1} = \boldsymbol{I} - \boldsymbol{V}_r \boldsymbol{D}_r \boldsymbol{V}_r^T + \mathcal{O}\left(\sum_{i=r+1}^{n} \frac{\lambda_i}{\lambda_i + 1} \right), \tag{4.9}$$

where $\boldsymbol{\Lambda}_r \in \mathbb{R}^{r \times r}$ and $\boldsymbol{V}_r \in \mathbb{R}^{n \times r}$ denote the truncated eigenvalue and eigenvector matrices, and $\boldsymbol{D}_r \overset{\text{def}}{=} \operatorname{diag}(\lambda_i/(\lambda_i + 1))$ where $\boldsymbol{D}_r \in \mathbb{R}^{r \times r}$. With this low-rank approximation, the final expression for approximation of the posterior covariance $\mathbf{\Gamma}_{\text{post}}$ is therefore given by

$$\mathbf{\Gamma}_{\text{post}} \approx \mathbf{\Gamma}_{\text{prior}} - \mathbf{\Gamma}_{\text{prior}}^{1/2} \boldsymbol{V}_r \boldsymbol{D}_r \boldsymbol{V}_r^T \mathbf{\Gamma}_{\text{prior}}^{1/2}. \tag{4.10}$$

The Frobenius norm of the error in the low-rank approximation of (4.9) is given by

$$\left\| \left(\tilde{\boldsymbol{H}}_{\text{misfit}} + \boldsymbol{I} \right)^{-1} - \left(\boldsymbol{I} - \boldsymbol{V}_r \boldsymbol{D}_r \boldsymbol{V}_r^T \right) \right\|_F = \sqrt{\sum_{j=r+1}^{n} \left(\frac{\lambda_j}{\lambda_j + 1} \right)^2} < \sqrt{\sum_{j=r+1}^{n} \lambda_j^2}, \tag{4.11}$$

and is in fact less than the approximation error for $\tilde{\boldsymbol{H}}_{\text{misfit}}$, and thus is also bounded independent of problem dimension. Finally, to relate the effect of the low-rank approximation of $\tilde{\boldsymbol{H}}_{\text{misfit}}$ on the posterior covariance $\mathbf{\Gamma}_{\text{post}}$, one would need to know the spectral structure of the prior covariance $\mathbf{\Gamma}_{\text{prior}}$, as is evident from (4.7).

This also provides us with a simple approximation to $\boldsymbol{H}^{-1/2}$ and its determinant, which is useful for computing samples of a local Gaussian approximation.

$$\begin{aligned} \boldsymbol{H}^{-1/2} &\approx \mathbf{\Gamma}_{\text{prior}}^{1/2} (\boldsymbol{V}_r \boldsymbol{P}_r \boldsymbol{V}_r^T + \boldsymbol{I}) \\ \det(\boldsymbol{H}^{1/2}) &\approx \det(\mathbf{\Gamma}_{\text{prior}}^{1/2}) \prod_{i=1}^{r} \sqrt{\lambda_i + 1} \end{aligned}$$

where $\boldsymbol{P}$ is a diagonal matrix with $\boldsymbol{P}_{jj} = \frac{1}{\sqrt{\lambda_j + 1}} - 1$.

Costs of construction of the low-rank approximation were presented in §4.1. Once the low-rank approximation has been constructed, the remaining costs to probe the posterior parameter density are linear in n. The costs are summarized in Table 4.1.

Calculation	cost for general $\mathbf{\Gamma}_{\text{prior}}$	cost for $\mathbf{\Gamma}_{\text{prior}} \propto \boldsymbol{I}$
variance: $\text{diag}(\mathbf{\Gamma}_{\text{post}})$	$(3r+1)n + r\gamma$	$(3r+1)n$
row/column: $\mathbf{\Gamma}_{\text{post}}\boldsymbol{e}_j$	$4rn + \gamma$	$2rn + r + 1$
multiplication: $\mathbf{\Gamma}_{\text{post}}\boldsymbol{v}$	$4rn + 2\gamma$	$(4r+1)n$
sample: $\exp(\boldsymbol{v}^T\mathbf{\Gamma}_{\text{post}}\boldsymbol{v})$	$(4r+2)n + 2\gamma$	$(4r+3)n$

Table 4.1: Cost of common computations using $\mathbf{\Gamma}_{\text{post}}$ once eigenvectors of $\tilde{\boldsymbol{H}}_{\text{misfit}}$, the prior-preconditioned Hessian of the data misfit, have been computed, where n is the dimension of $\boldsymbol{x}$, r is the dimension of the low-rank approximation of $\tilde{\boldsymbol{H}}_{\text{misfit}}$, and γ is the cost of multiplying $\mathbf{\Gamma}_{\text{prior}}^{1/2}$ by a vector. The last column corresponds to the case of an i.i.d. prior.

4.3 Analysis of steady state groundwater flow inverse problem

We now turn to the nonlinear inverse problem in groundwater flow introduced in Chapter 3. The section will examine a 1D infinite-dimensional steady state flow inverse problem to investigate the suitability of a low-rank approximation to the Hessian, then present numerical results for a 2D example to support its use for high-dimensional parameter spaces.

4.3.1 Analysis of a model inverse problem, 1D

The forward problem, expressing the relationship of the head h and the log hydraulic conductivity γ over the domain $(0, L)$, is

$$\begin{aligned} -\frac{d}{dx}\left(e^{\gamma}\frac{dh}{dx}\right) &= 0, \qquad (4.12)\\ h(0) &= h_0,\\ h(L) &= h_L, \end{aligned}$$

where h_0 and h_L are Dirichlet boundary conditions. The general form of the Hessian, adjoint equation, incremental adjoint equation, and incremental state equation are derived in §3.2 through calculus of variations. The system of ODEs in this section are simplified forms of those results.

We consider two observation operators - one in which the head is measured over the entire domain, and one with n point measurements at $x_j = \frac{L}{n}(j - 1/2)$ for $j = 1, \ldots, n$. Mathematically we define the observation operators as

$$b_{\text{noise}} = \begin{cases} \mathbf{1} & \text{full observations} \\ \frac{L}{n}\sum_{j=1}^{n}\delta(x - x_j) \text{ for } j = 1, \ldots, n & \text{point observations} \end{cases}$$

and the corresponding objective functions are the infinite dimensional functionals

$$J^{\infty}(h, \gamma) = \frac{1}{2}\int_0^L (h(x) - h_{\text{obs}}(x))^2 dx$$

and

$$J^{n}(h, \gamma) = \frac{L}{2n}\sum_{j=1}^{n}(h(x_j) - h_{\text{obs}}(x_j))^2$$

where $h(x)$ satisfies (4.12).

We assume zero noise in the observations, and that we are evaluating the Hessian at a point such that our observations match the state: $h(x) = h_{\text{obs}}$. The adjoint equation is therefore

$$\begin{aligned}\frac{d}{dx}\left(e^{\gamma}\frac{dq}{dx}\right) &= 0\\ q(0) &= 0\\ q(L) &= 0\end{aligned}$$

For simplicity, we will focus on the eigenpairs of the Hessian at $\gamma = \gamma_0$ constant. The solution to the state and adjoint equations are then

$$\begin{aligned}&h(x) = h_0 + \left(\frac{h_L - h_0}{L}\right)x,\\ &q(x) = 0.\end{aligned}$$

Incorporating these solutions into the incremental state and adjoint equations, we obtain

Incremental state	Incremental adjoint
$-\frac{d^2\tilde{h}}{dx^2} = \left(\frac{h_L - h_0}{L}\right)\frac{d\tilde{\gamma}}{dx},$	$e^{\gamma_0}\frac{d^2\tilde{q}}{dx^2} = \tilde{h}\, b_{\text{noise}},$
$\tilde{h}(0) = 0,$	$\tilde{q}(0) = 0,$
$\tilde{h}(L) = 0,$	$\tilde{q}(L) = 0.$

Finally, the eigenvalue problem for the Hessian, its eigenfunctions $\tilde{\gamma}$, and its eigenvalues λ, is

$$H(\gamma_0, \tilde{\gamma}) = e^{\gamma_0}\frac{dh}{dx}\frac{d\tilde{q}}{dx} = \left(\frac{e^{\gamma_0}(h_L - h_0)}{L}\right)\frac{d\tilde{q}}{dx} = \lambda\tilde{\gamma}.$$

We begin with the case of full observations over the domain: the mth eigenpair can be shown to be the expressions

$$\tilde{\gamma}_m = \sqrt{\frac{2}{L}} \cos\left(\frac{m\pi x}{L}\right) \qquad\qquad \lambda_m = \frac{(h_L - h_0)^2}{\pi^2 m^2} \tag{4.13}$$

for $m = 1, \ldots, \infty$. For completeness, the expressions for the incremental state and adjoint are

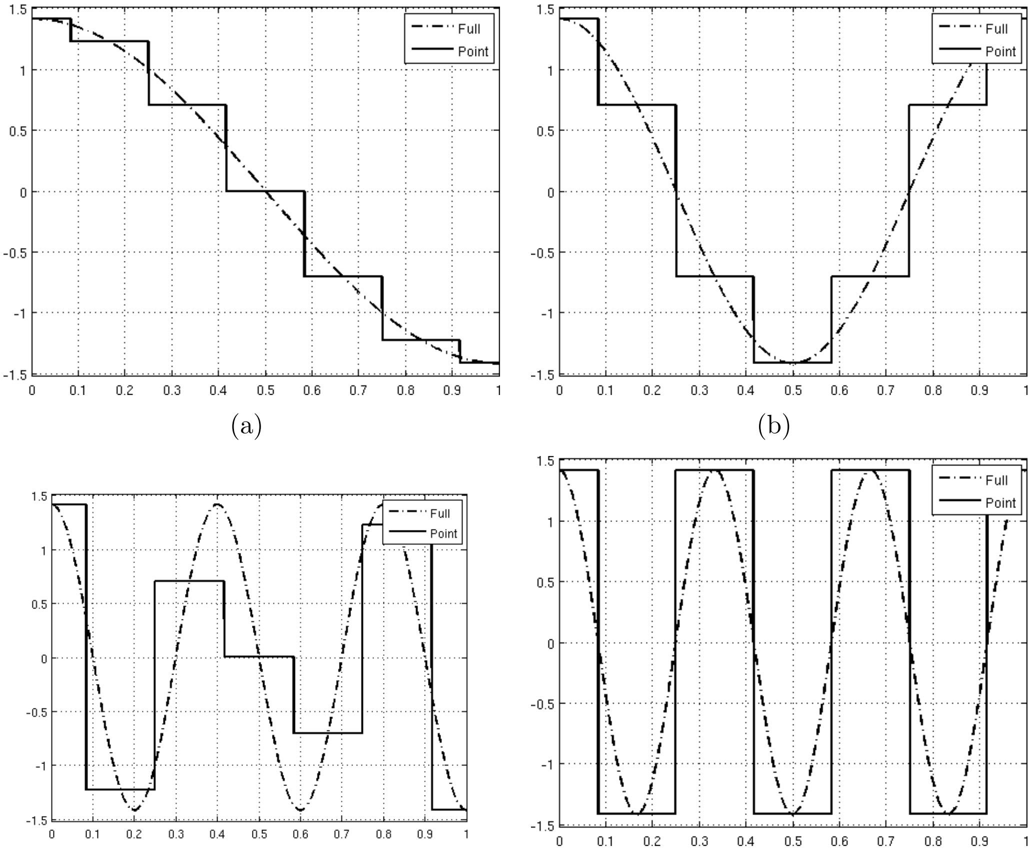
$$\begin{aligned} \tilde{h}_m &= -\sqrt{\frac{2}{L}} \left(\frac{h_L - h_0}{m\pi}\right) \sin\left(\frac{m\pi x}{L}\right), \\ \tilde{q}_m &= \sqrt{\frac{2}{L}} \left(\frac{(h_L - h_0)L^2}{e^{\gamma_0} m^3 \pi^3}\right) \sin\left(\frac{m\pi x}{L}\right). \end{aligned}$$

We now turn to the case of n point observations. The eigenfunctions are piecewise constant interpolations of cosine waves matching the midpoint between observation points, with jumps in value at the observation points. There are n eigenpairs (consistent with the Nyquist sampling theorem for reconstruction of waves based on discrete sample points). For notational convenience, we define $x_0 = 0$ and $x_{n+1} = L$, the left and right endpoints of the domain. The mth eigenpair is described by the expressions

$$\begin{aligned} \tilde{\gamma}_m(x) &= \beta \cos\left(\frac{m\pi j}{n}\right) \text{ for } x_j \le x \le x_{j+1} \text{ and } j = 0, \ldots, n+1 \\ \lambda_m &= \frac{(h_L - h_0)^2}{4n^2 \sin(m\pi/2n)^2} \end{aligned} \tag{4.14}$$

where the normalization constant

$$\beta_m = \begin{cases} 2\left(1 + 2n + \frac{\sin(m(-1+2n)\pi/n)}{\sin(m\pi/n)}\right)^{-1/2} & \text{if } 1 \le m < n \\ 1/\sqrt{n} & \text{if m=n.} \end{cases}$$



(a) (b)

(c) (d)

Figure 4.1: Four eigenfunctions of the Hessian at $\gamma = \gamma_0$ for the full observation operator and the pointwise observation operator with 6 evenly spaced points. See equations 4.13 and 4.14 for the expressions. Note that the eigenfunctions associated with the pointwise observation operator are piecewise constant interpolations of the eigenfunctions associated with the full observation operator. (a) 1st eigenfunctions. (b) 2nd eigenfunctions. (c) 5th eigenfunctions. (d) 6th eigenfunctions.

The eigenvalues are zero for $m > n$. The eigenfunctions for both cases are illustrated in Figure 4.1.

We define a piecewise linear function $f(x_j)$ such that $f(x_j) = \sin\left(\frac{m\pi x_j}{L}\right)$ for $j = 0, \ldots, n+1$. Let

$$f(x) = \begin{cases} a_0 x & (0, x_1) \\ a_1 x + b_1 & (x_1, x_2) \\ \vdots & \vdots \\ a_{n-1} x + b_{n-1} & (x_{n-1}, x_n) \\ a_n x - a_n & (x_n, 1) \end{cases}$$

where

$$\begin{aligned} a_j(x) &= \frac{1}{x_{j+1} - x_j}\left(\sin\left(\frac{m\pi x_{j+1}}{L}\right) - \sin\left(\frac{m\pi x_j}{L}\right)\right) \\ b_j(x) &= \sum_{k=1}^{j}(a_{k-1}x_k - a_k x_k) \end{aligned}$$

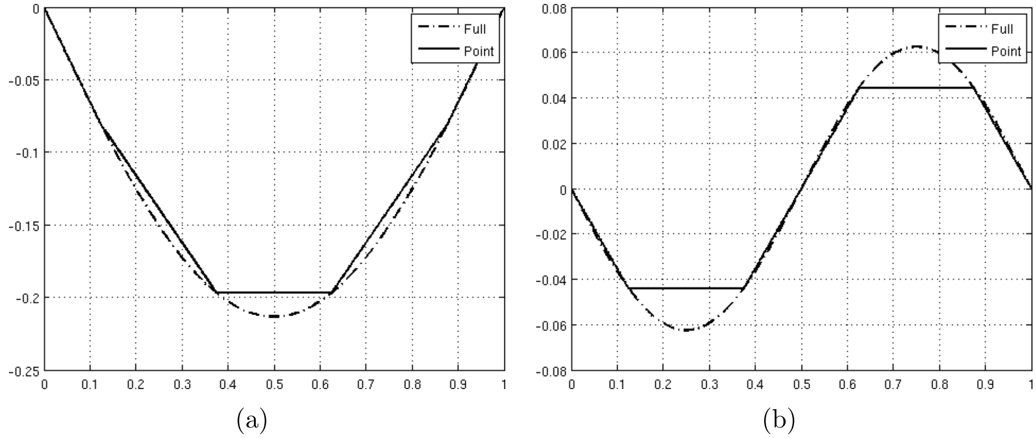
for $j = 0, \ldots, n$.

Then the expressions for the incremental state and adjoint are piecewise linear interpolations of sine waves given by

$$\begin{aligned} \tilde{h}_m(x) &= -\frac{(h_L - h_0)\beta}{2n\sin(m\pi/2n)} f(x) \\ \tilde{q}_m(x) &= \frac{(h_L - h_0)L^2\beta}{8e^{\gamma_0} n^3 \sin(m\pi/2n)^3} f(x), \end{aligned}$$

as illustrated in Figure 4.2.

We continue with analysis of the eigenpairs of the Hessian for the two observation operators. First note that $J^n(h, \gamma)$ is equivalent to a Riemann sum approximation of $J^\infty(h, \gamma)$ with the midpoint approximation. Thus we



(a) (b)

Figure 4.2: Plots of the incremental state solution for the full observation operator and the 4 point observation operator. (a) $m = 1$. (b) $m = 2$.

should expect that as $n \to \infty$, the eigenpairs of the Hessian should also converge. We can see that the eigenfunctions converge in Figure 4.3 (in fact, note their resemblance to Riemann sums). Larger eigenvalues are associated with smoother eigenfunctions. As λ decreases, the eigenfunctions are increasingly oscillatory.

Physically, the eigenvalues depend on the pressure difference across the domain $(h_L - h_0)$. The eigenvalues for the full observation case decay quadratically, inviting a low rank approximation. In addition,

$$\begin{aligned}\lambda_m^n - \lambda_m^{\text{full}} &= \frac{(h_L - h_0)^2}{4n^2}\left(\frac{1}{\sin(m\pi/2n)^2} - \frac{4n^2}{\pi^2 m^2}\right)\\ &= \frac{(h_L - h_0)^2}{12}\left(\frac{1}{n^2} + \frac{\pi^2}{20}\frac{m^2}{n^4} + O\left(\frac{m^4}{n^6}\right)\right),\end{aligned}$$

so we see the eigenvalues based on the n-point observation operator converge quadratically to the eigenvalues based on the full observation operator. Smaller values of m converge faster. This effect is visible in Figure 4.4. Thus, although we need no more than n eigenpairs for n observations, a low rank approximation may still be appropriate for large n. For example, in the graph of error versus number of retained eigenpairs in Figure 4.5(c), the error drops exponentially and indicates a low rank approximation is appropriate well before n points.

For a given cutoff $\lambda > \alpha$, the number of retained eigenvalues r for the full observation operator is

$$r = \left\lfloor \frac{|h_L - h_0|}{\pi\sqrt{\alpha}} \right\rfloor$$

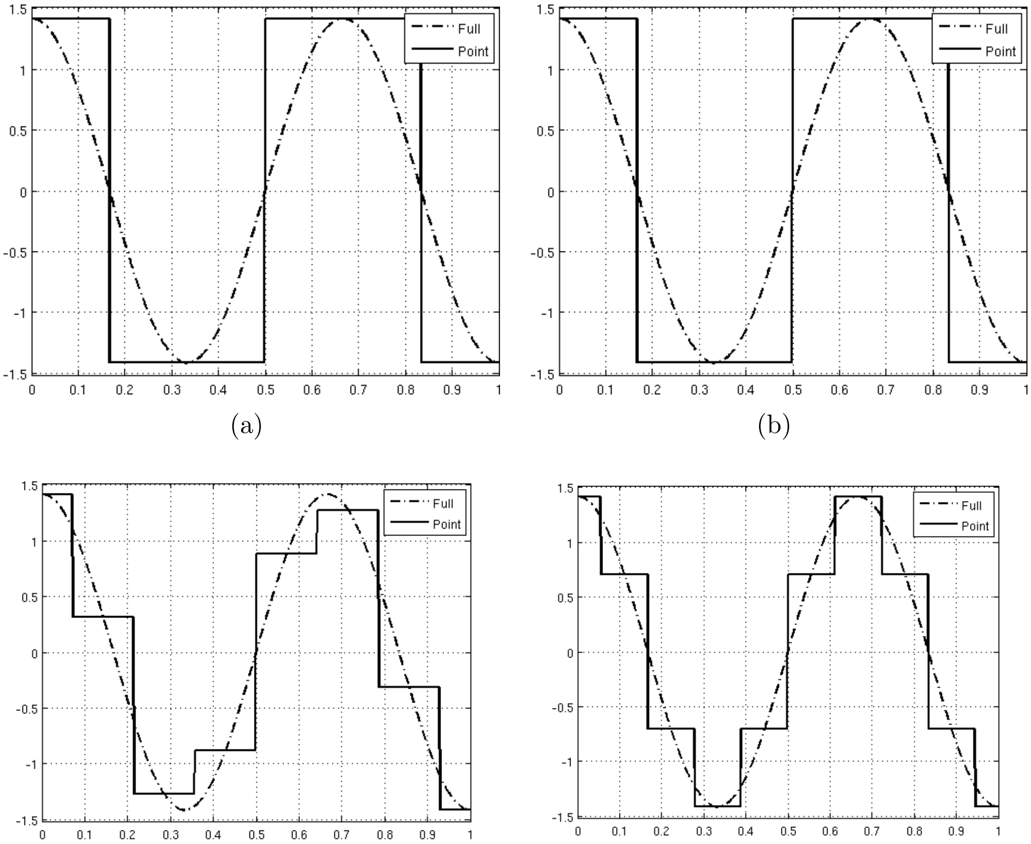
and for n point observations is

$$r = \begin{cases} \left\lfloor \frac{2n}{\pi} \arcsin\left(\frac{|h_L - h_0|}{2n\sqrt{\alpha}}\right) \right\rfloor & \alpha \geq \frac{(h_L - h_0)^2}{4n^2} \\ n & \text{otherwise} \end{cases}$$

We next analyze the error in the posterior covariance matrix due to the choice of truncation r. Assuming the eigenvalues of the finite dimensional posterior decay with the infinite dimensional expressions, the squared error in the Frobenius norm is

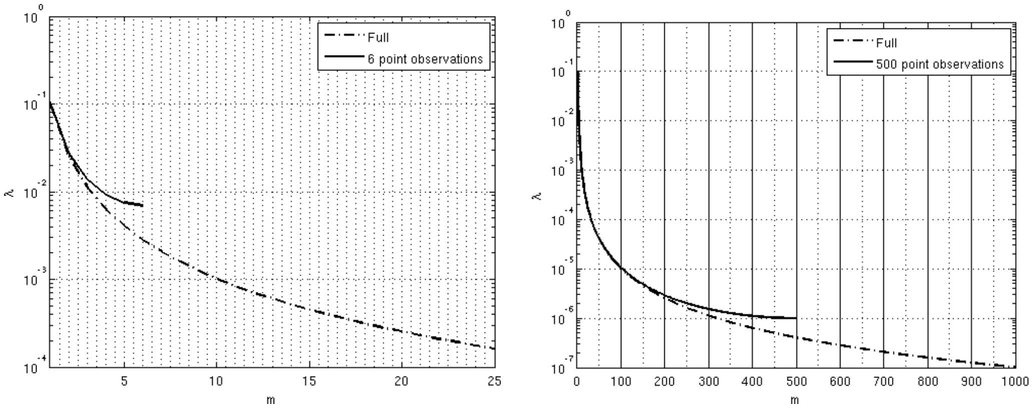
$$\sum_{j=r+1}^{\infty} \lambda_j^2 = \frac{(h_L - h_0)^4}{\pi^4} \sum_{j=r+1}^{\infty} \frac{1}{j^4} = \frac{(h_L - h_0)^4}{6\pi^4} \psi^{(3)}(r+1)$$

for the full observation operator where $\psi^{(3)}(x)$ is the polygamma function of order 3 (i.e. the 4th derivative of the logarithm of the gamma function). For the n point observation operator, we derive an upper bound by treating the



(a) (b) (c) (d)

Figure 4.3: The third eigenfunctions of the Hessian at $\gamma = \gamma_0$ for the full observation operator and the pointwise observation operator. See equations 4.14 and 4.14 for the expressions. (a) 3 points. (b) 4 points. (c) 7 points. (d) 9 points.



(a) (b)

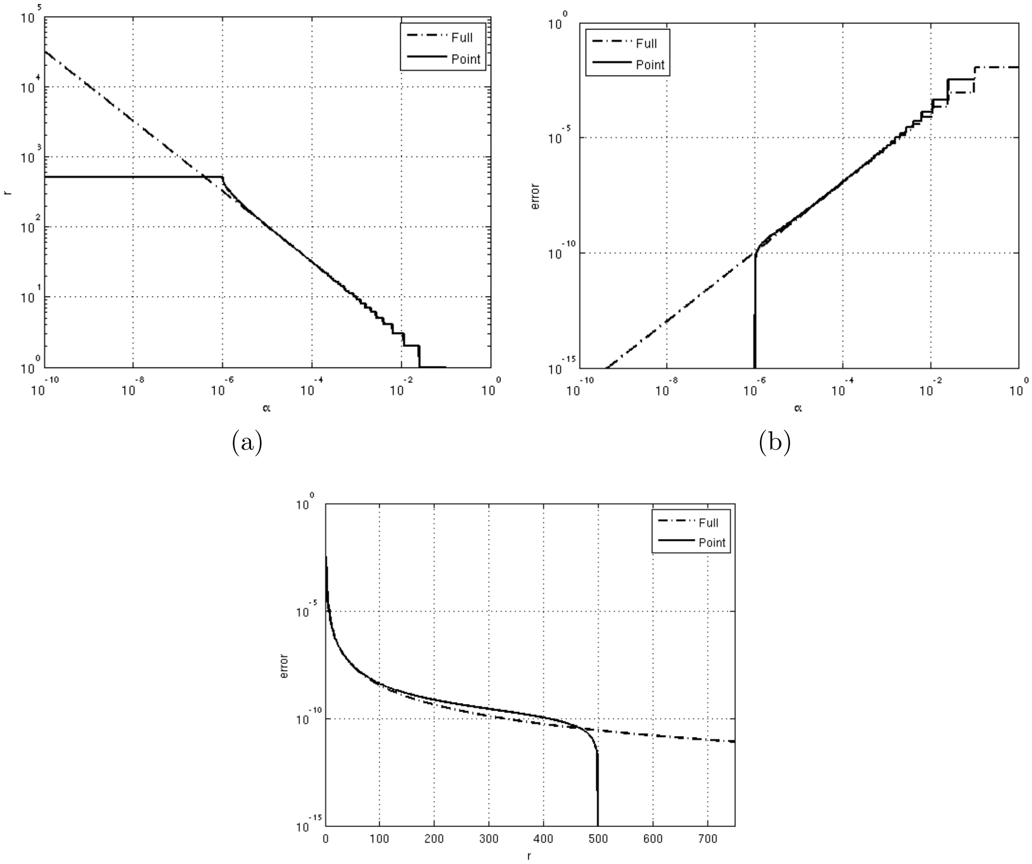
Figure 4.4: Spectrum of the Hessian at $\gamma = \gamma_0$ for the full observation operator and the pointwise observation operator. See equations 4.13 and 4.14 for the expressions. (a) 6 points. (b) 500 points.

sum as a right Riemann sum, so that

$$\begin{aligned}\sum_{j=r+1}^{n} \lambda_j^2 &< \int_r^n \frac{(h_L - h_0)^4}{16n^4 \sin(\pi x/2n)^4} \\ &< \frac{(h_L - h_0)^4}{24n^3\pi} \left(2 - \cos\left(\frac{\pi r}{n}\right)\right) \cot\left(\frac{\pi r}{2n}\right) \csc\left(\frac{\pi r}{2n}\right)^2 .\end{aligned}$$

A comparison of the bounds based on the two observation operators is given in Figure 4.5.

We finish the section with some additional calculations of derivatives (using trigonometric angle sum formulas) to help the reader verify the preced-



(a)

(b)

(c)

Figure 4.5: Plots show comparisons between the continuous observation operator and the 500 point observation operator. (a) Dependence of the number of retained eigenvalues r on a given cutoff $\lambda > \alpha$. (b) Estimate of upper bound on the squared Frobenius error in the low-rank approximation based on the choice of a cutoff value α. (c) Relationship between the number of retained eigenvalues r and the upper bound on the error in the low rank approximation.

ing expressions.

$$\begin{aligned}
\frac{d\tilde{\gamma}}{dx} &= \beta \sum_{j=1}^{n} \left(\cos\left(\frac{m\pi j}{n}\right) - \cos\left(\frac{m\pi(j-1)}{n}\right)\right) \delta(x-x_j) \\
&= \beta \sum_{j=1}^{n} \left(\cos\left(\frac{m\pi(j-1/2)}{n} + \frac{m\pi}{2n}\right) - \cos\left(\frac{m\pi(j-1/2)}{n} - \frac{m\pi}{2n}\right)\right) \delta(x-x_j) \\
&= \beta \sum_{j=1}^{n} \left(\cos\left(\frac{m\pi(j-1/2)}{n}\right) \cos\left(\frac{m\pi}{2n}\right) - \sin\left(\frac{m\pi(j-1/2)}{n}\right) \sin\left(\frac{m\pi}{2n}\right) \right. \\
&\left. - \cos\left(\frac{m\pi(j-1/2)}{n}\right) \cos\left(\frac{m\pi}{2n}\right) - \sin\left(\frac{m\pi(j-1/2)}{n}\right) \sin\left(\frac{m\pi}{2n}\right)\right) \delta(x-x_j) \\
&= -2\beta \sum_{j=1}^{n} \sin\left(\frac{m\pi(j-1/2)}{n}\right) \sin\left(\frac{m\pi}{2n}\right) \delta(x-x_j).
\end{aligned}$$

For $x_j \leq x \leq x_{j+1}$ and $1 < j < n$,

$$\begin{aligned}
\frac{df}{dx} &= \frac{n}{L}\left(\sin\left(\frac{m\pi(j+1/2)}{n}\right) - \sin\left(\frac{m\pi(j-1/2)}{n}\right)\right) \\
&= \frac{h_1 n}{e^{\gamma_0} L}\left(\sin\left(\frac{m\pi j}{n} + \frac{m\pi}{2n}\right) - \sin\left(\frac{m\pi j}{n} - \frac{m\pi}{2n}\right)\right) \\
&= \frac{n}{L}\left(\sin\left(\frac{m\pi j}{n}\right) \cos\left(\frac{m\pi}{2n}\right) + \cos\left(\frac{m\pi j}{n}\right) \sin\left(\frac{m\pi}{2n}\right) \right. \\
&\left. - \sin\left(\frac{m\pi j}{n}\right) \cos\left(\frac{m\pi}{2n}\right) + \cos\left(\frac{m\pi j}{n}\right) \sin\left(\frac{m\pi}{2n}\right)\right) \\
&= \frac{2n}{L} \cos\left(\frac{m\pi j}{n}\right) \sin\left(\frac{m\pi}{2n}\right).
\end{aligned}$$

For $0 \leq x \leq x_1$,

$$\frac{df}{dx} = \frac{2n}{L}\left(\sin\left(\frac{m\pi}{2n}\right) - 0\right) = \frac{2n}{L} \sin\left(\frac{m\pi}{2n}\right) \cos\left(\frac{m\pi j}{n}\right) \text{ for } j = 0.$$

For $x_n \leq x \leq 1$

$$\frac{df}{dx} = \frac{2n}{L}\left(0 - \sin\left(\frac{m\pi(n-\frac{1}{2})}{n}\right)\right) = \frac{2n}{L} \sin\left(\frac{m\pi}{2n}\right) \cos\left(\frac{m\pi j}{n}\right) \text{ for } j = n.$$

For $x = x_j$ and $j = 1, \ldots, n$

$$\begin{aligned}
\frac{d^2 f}{dx^2} &= \frac{2n}{L} \sin\left(\frac{m\pi}{2n}\right) \left(\cos\left(\frac{m\pi j}{n}\right) - \cos\left(\frac{m\pi(j-1)}{n}\right)\right) \delta(x - x_j) \\
&= \frac{2n}{L} \sin\left(\frac{m\pi}{2n}\right) \left(\cos\left(\frac{m\pi(j-1/2)}{n} + \frac{m\pi}{2n}\right)\right. \\
&\left. - \cos\left(\frac{m\pi(j-1/2)}{n} - \frac{m\pi}{2n}\right)\right) \delta(x - x_j) \\
&= \frac{2n}{L} \sin\left(\frac{m\pi}{2n}\right) \left(\cos\left(\frac{m\pi(j-1/2)}{n}\right) \cos\left(\frac{m\pi}{2n}\right)\right. \\
&- \sin\left(\frac{m\pi(j-1/2)}{n}\right) \sin\left(\frac{m\pi}{2n}\right) - \cos\left(\frac{m\pi(j-1/2)}{n}\right) \cos\left(\frac{m\pi}{2n}\right) \\
&\left. - \sin\left(\frac{m\pi(j-1/2)}{n}\right) \sin\left(\frac{m\pi}{2n}\right)\right) \delta(x - x_j) \\
&= -\frac{4n}{L} \sin\left(\frac{m\pi}{2n}\right)^2 \sin\left(\frac{m\pi(j-1/2)}{n}\right) \delta(x - x_j).
\end{aligned}$$

4.3.2 Numerical experiments

We introduce a nonlinear statistical inverse problem governed by the steady state form of the equations in Chapter 3 for a two-dimensional groundwater example. The forward problem is described by the PDE with boundary conditions

$$\begin{aligned}
-\nabla \cdot e^{\gamma} \nabla h &= f & &(0, 6\text{km}) \times (0, 6\text{km}) \\
h(x, 0) &= 100\text{m} & -e^{\gamma} \nabla h \cdot \boldsymbol{n}(6\text{km}, y) &= 0 \\
-e^{\gamma} \nabla h \cdot \boldsymbol{n}(0, y) &= 500\text{m}^3\text{km}^{-1}\text{day}^{-1} & -e^{\gamma} \nabla h \cdot \boldsymbol{n}(x, 6\text{km}) &= 0
\end{aligned}$$

where the recharge term is

$$f(x,y) = \begin{cases} 0 & 0 < y < 4\text{km} \\ 0.137 \times 10^{-3}\text{m/day} & 4\text{km} < y < 5\text{km} \\ 0.274 \times 10^{-3}\text{m/day} & 5\text{km} < y < 6\text{km} \end{cases}.$$

The boundary conditions, domain, and recharge term are extracted from the example in [36] and [31].

We generate a synthetic log transmissivity field (see Figure 4.6(a)) through sampling a Gaussian process with mean $\mu = 3$, whose covariance is defined by the Mátern covariance function,

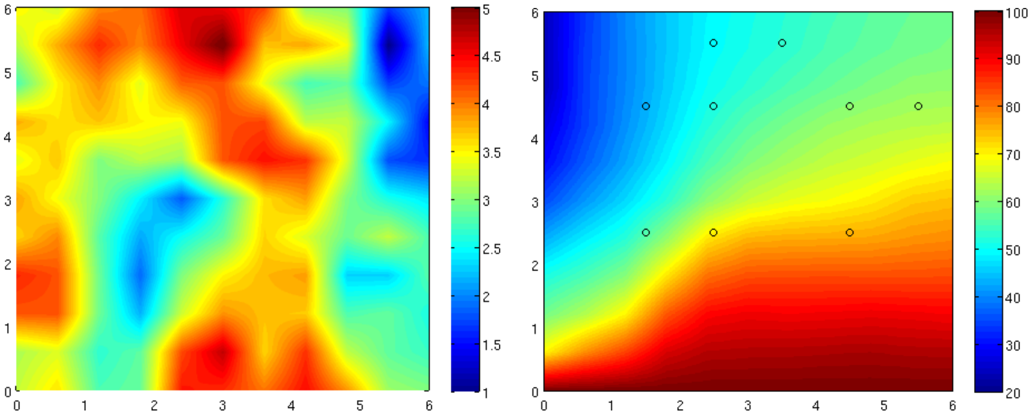
$$C(r) = \sigma^2 2^{1-\nu} \left(\frac{r}{\theta}\right)^{\nu} \frac{\mathcal{K}_\nu(r/\theta)}{\Gamma(\nu)}, \tag{4.15}$$

where Γ is the gamma function and $\mathcal{K}_\nu$ is the modified Bessel function of the second kind, and $\theta = 1$, $\nu = 1$, and $\sigma = 1$. The Mátern covariance depends on r, the distance between points and has a parameter, ν, that determines smoothness of the field. It is often used to generate synthetic permeability fields.

In addition, we define a pointwise observation operator (in general, this must be defined with a mollifier to ensure well-posedness of the adjoint problem in the limit as the mesh size goes to zero)

$$b(\boldsymbol{x}) = \sum_{j=1}^{9} \exp\left(-\frac{1}{D - \|\boldsymbol{x} - \boldsymbol{x}_j\|_2^2}\right),$$

where D is the width of the observation well and with the points $\boldsymbol{x}_j$ given by Table 4.2. The observation points are illustrated in Figure 4.6. For the example



(a) (b)

Figure 4.6: (a) The “ground truth” log transmissivity: synthetic field generated as a sample of a Gaussian process with mean $\mu = 3$ and Mátern covariance from Eqn 4.15 with parameters $\theta = 1$, $\nu = 1$, and $\sigma = 1$. (b) Piezometric head for the field in (a). The black circles mark observation points. See Table 4.2 for observation values for an example with 121 degrees of freedom.

on a coarse mesh in our results section, we use the pointwise observation operator

$$b(\boldsymbol{x}) = \sum_{j=1}^{9} \delta(x - x_j)$$

without a mollifier. To lessen the inverse crime, we add 1% noise to the observed head to generate the synthetic data for the statistical inverse problem.

We continue to the Bayesian formulation of a statistical inverse problem with Gaussian prior and noise (see §2.1). We define the noise pdf as a Gaussian such that

$$\pi_{\text{noise}}(\boldsymbol{e}) = \mathcal{N}(0, \alpha \boldsymbol{I})$$

where α is a constant defined in Table 4.2, consistent with our synthetic ob-

j	$\boldsymbol{x}_j$	Exact observed head	Synthetic noise	Total
0	$(2.5, 5.5)$	49.45082	-0.06493812	49.38588
1	$(3.5, 5.5)$	52.70623	0.4225783	53.12881
2	$(1.5, 4.5)$	43.15628	0.9965889	44.15287
3	$(2.5, 4.5)$	51.70564	0.1224217	51.82806
4	$(4.5, 4.5)$	60.20506	-0.2248143	59.98025
5	$(5.5, 4.5)$	62.16391	0.4810087	62.64492
6	$(1.5, 2.5)$	55.52143	0.8825147	56.40394
7	$(2.5, 2.5)$	70.20701	-0.3307525	69.87625
8	$(4.5, 2.5)$	75.29253	0.1424353	75.43496

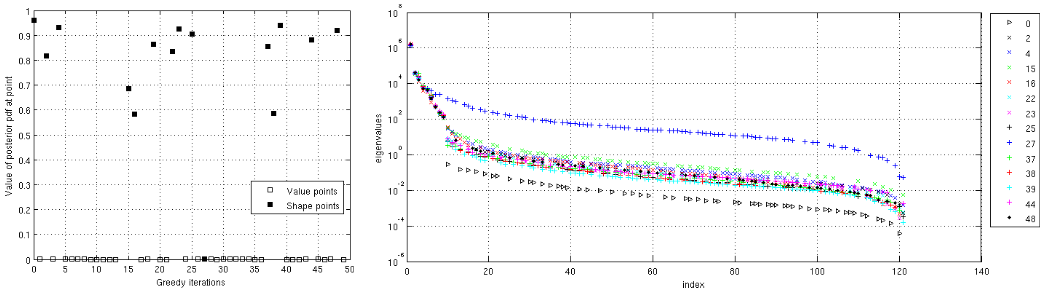
Table 4.2: Observations of the piezometric head plus noise. Noise generated from an iid normal distribution with mean 0 and standard deviation 0.4420484, so that the variance α of noise distribution is set to 1% noise as calculated through the expression $0.01\sqrt{\sum_{j=1}^{9} h(x_j)^2/9}$.

servations. The prior pdf is a Gaussian

$$\pi_{\text{prior}} = \mathcal{N}(\bar{\gamma}, \boldsymbol{\Gamma}_{\text{prior}})$$

where the prior mean is constant such that $\bar{\gamma} = 3.290624$, the average value of the synthetic control field in Figure 4.6. The prior covariance is defined as described in Chapter 3, with parameters 0.05 and 0.03.

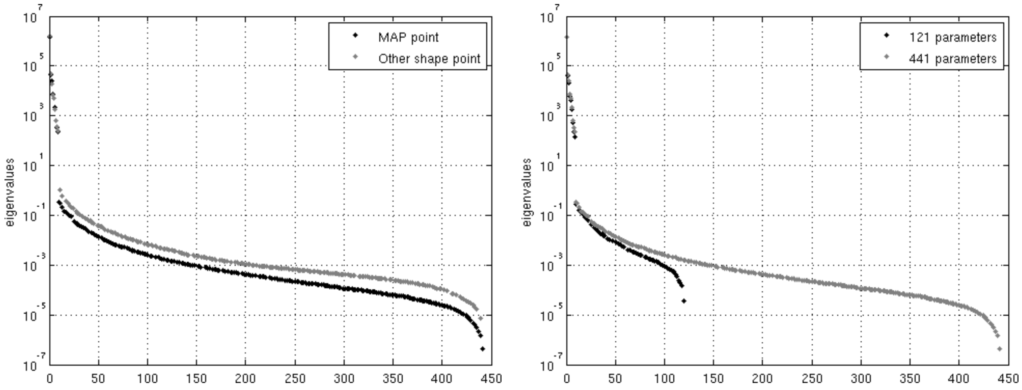
The data for this figure was developed by running the greedy algorithm described in §2.2 for 50 iterations. The black squares in Figure 4.7(a) represent points at which a low-rank approximation was built (plotted according to the value at that point of the non-normalized posterior pdf). The positive portion of the spectrum at each point is graphed in Figure 4.7(b) on a log scale. Note that all display the decay necessary to support a low-rank approximation. There are two outliers - the smallest corresponds to the MAP point, and the largest corresponds to the one shape point in Figure 4.7(a) with an unusually small value. Another feature, seen even more clearly in Figure 4.8, is that the first nine eigenvalues (corresponding to the nine observations) are considerably larger than the rest of the spectrum.



(a) (b)

Figure 4.7: (a) Values of the (non-normalized) posterior pdf at the points found in the optimization stage of the greedy algorithm for fifty iterations. Solid squares mark the shape points, at which we form a low-rank approximation to the Hessian. (b) Positive eigenvalues of the prior-preconditioned Hessian of the data misfit at the shape points. Note the lowest, "Shape point 0" is the MAP point, and the highest, "Shape point 27" had an unusually small value.

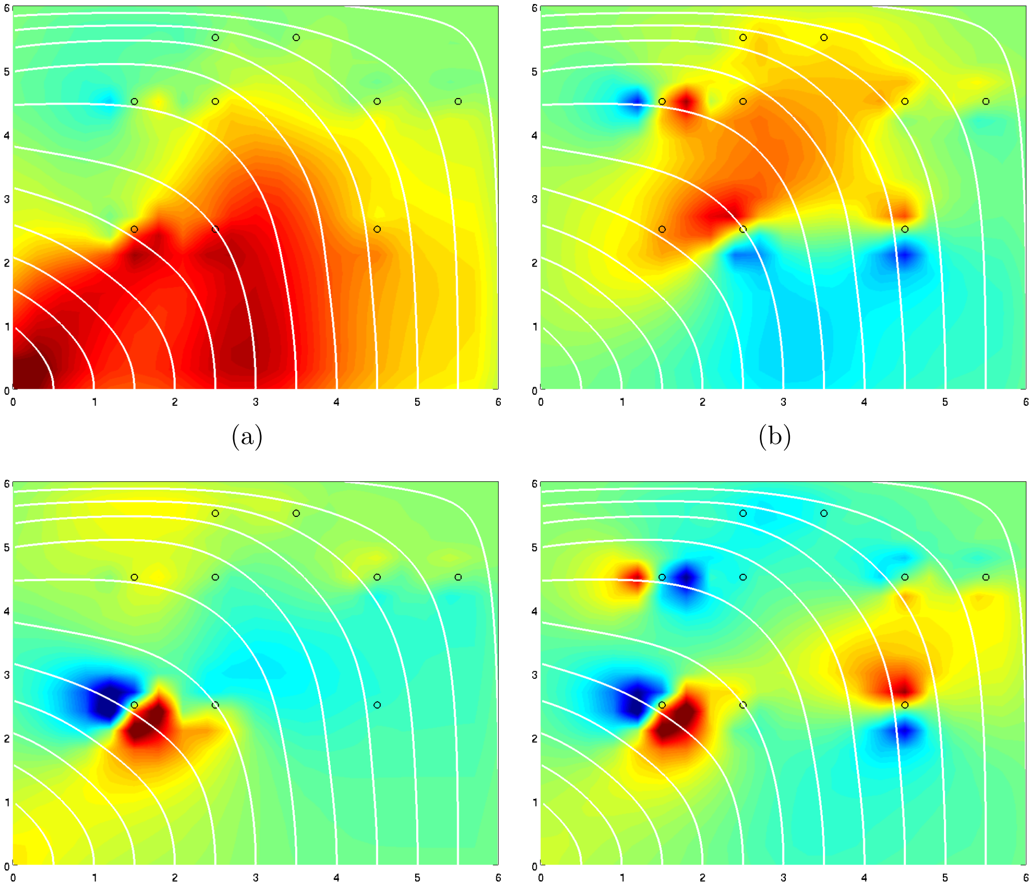
To investigate the scalability of the low-rank Hessian approximation, we turn to Figure 4.8. We reproduce the problem described for 121 parameters with 441 parameters. Figure 4.8(a) demonstrates the viability of the low rank approximation at two points chosen through the greedy algorithm, showing the same structure as in Figure 4.7(b). Then, in Figure 4.8(b), we compare the spectra at the MAP point for the two cases directly. Over a range of 7 orders of magnitude, the two meshes have essentially the same eigenvalues, and hence the same data content. Both display a jump after 9 eigenvalues, which is consistent with the analysis of the 1D problem in §4.3.1.



(a) (b)

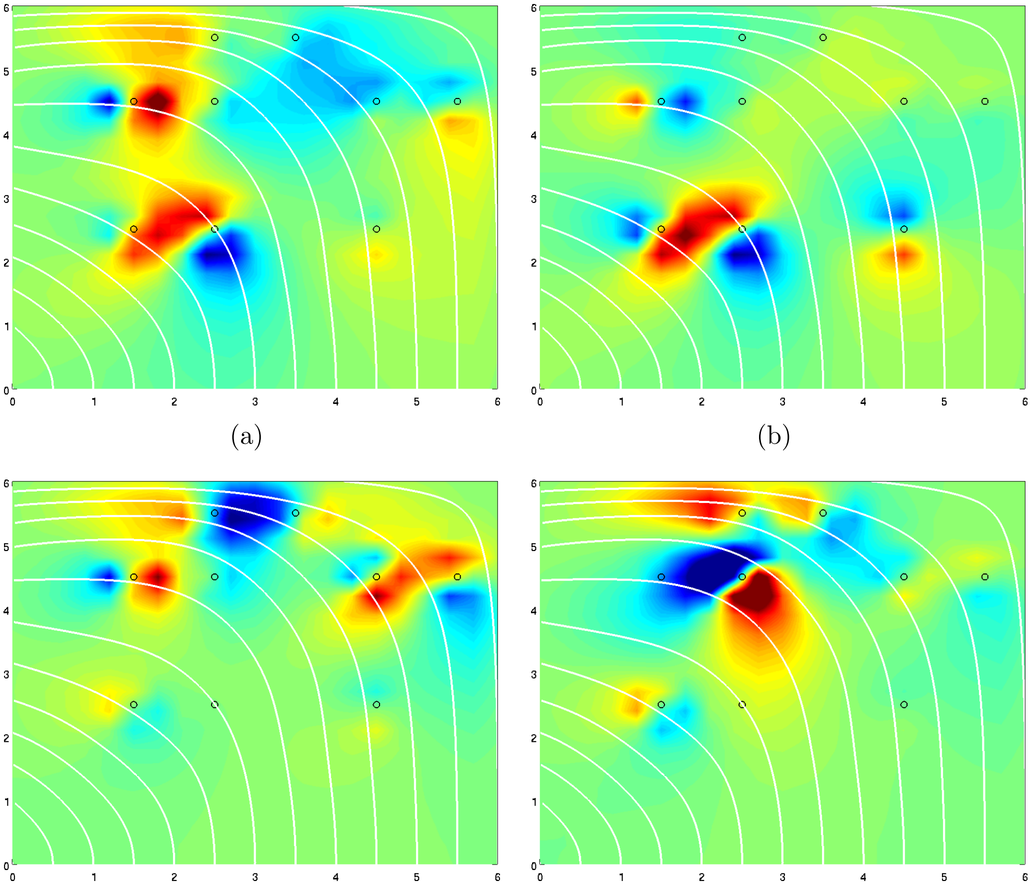
Figure 4.8: (a) Spectrum at the MAP point and an additional shape point from the 441 parameter version of the problem. (b) Comparison of the spectrum of the prior-preconditioned Hessian of the data misfit at the MAP point for 121 parameters vs 441 parameters.

We present the first nine eigenvectors at the MAP point (from the 441 parameter case, but corresponding to similar eigenvectors in the 121 parameter case) in Figures 4.9–4.11. Streamlines of the flow are in white, and black circles are sensor locations. In the figures you can see jumps at the observation points, where the orientation of the jump is determined by the direction of the streamlines (which align with the pressure gradient). The eigenvector associated with the largest eigenvalue is the smoothest, and later eigenvectors display increasingly oscillatory behavior.



(a) (b) (c) (d)

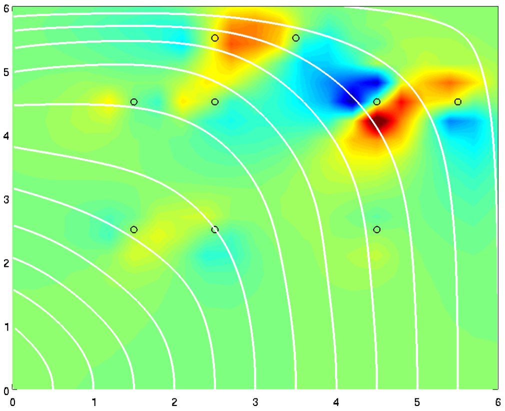
Figure 4.9: First through fourth eigenvectors of the Hessian of the data misfit at the MAP point, ordered by the magnitude of the associated eigenvalues of prior-preconditioned Hessian of the data misfit. Eigenvalues (a) 1,500,422 (b) 42,576.25 (c) 24,075.56 (d) 7,309.192. Streamlines of the flow are marked in white. Black circles mark observation points.



(a) (b)

(c) (d)

Figure 4.10: Fifth through eighth eigenvectors of the Hessian of the data misfit at the MAP point, as ordered by the magnitude of the associated eigenvalues of prior-preconditioned Hessian of the data misfit. Eigenvalues (a) 5,164.517 (b) 2,260.133 (c) 623.9096 (d) 328.8667. Streamlines of the flow are marked in white. Black circles mark observation points.



(a)

Figure 4.11: Ninth eigenvector of the Hessian of the data misfit at the MAP point, as ordered by the magnitude of the associated eigenvalues of prior-preconditioned Hessian of the data misfit. The associated eigenvalue 230.6024. Streamlines of the flow are marked in white. Black circles mark observation points.

Chapter 5

A 3D convection-diffusion inverse problem

We consider the case of a linear ill-posed inverse problem within the framework of Bayesian inference as a stepping stone to the nonlinear groundwater flow inverse problem in Chapter6. This chapter will thoroughly investigate the low-rank approximation to the Hessian for up to 1.4 million parameters. For a linear inverse problem, an additive Gaussian noise model and a Gaussian prior density of the model parameters will result in a Gaussian posterior pdf for the model parameters. Thus, the statistical solution of the inverse problem, i.e. the posterior pdf, can be completely described by its mean and covariance matrix. The mean is given by maximizing the posterior pdf, which leads to a linear least squares optimization problem whose structure is identical to that of a properly weighted deterministic inverse problem. Thus, state-of-the-art numerical algorithms from large-scale deterministic inverse problems can be exploited to find the mean. On the other hand, the posterior covariance matrix $\mathbf{\Gamma}_{\text{post}}$ is given by the inverse of the Hessian matrix of this least squares function. Due to the high-dimensionality of the parameter space and expense of solving the underlying forward and adjoint PDEs needed to construct the Hessian, direct computation of $\mathbf{\Gamma}_{\text{post}}$ is usually intractable. Moreover, the explicit inverse of the Hessian is not required for deterministic inversion, and thus one cannot appeal to deterministic techniques for directly computing $\mathbf{\Gamma}_{\text{post}}$. Instead, we employ the low-rank approximation described in Chapter 4 for the prior-preconditioned Hessian of the data misfit, then invert using the Sherman-Morrison-Woodbury formula.

We now introduce a large-scale statistical inverse problem governed by three-dimensional convective-diffusive contaminant transport. The inverse problem seeks to find the initial concentration field of the contaminant and associated uncertainty, given noisy measurements of contaminant concentration at certain spatio-temporal locations or at final time, prior information on the initial contaminant, and the PDE model of the transport of the contaminant. In this section we also present analysis of a related one-dimensional convection-diffusion inverse problem to provide insight into the numerical experiments on the 3D inverse problem in §5.3.

5.1 Problem description

The forward problem is a discretization of a convection-diffusion PDE describing the evolution of the contaminant:

$$\begin{aligned} u_t - k\triangle u + \boldsymbol{v}\cdot\nabla u &= 0, && \Omega\times(0,T),\\ u &= u_0, && \Omega\times\{t=0\}, \qquad (5.1)\\ k\nabla u\cdot\boldsymbol{n} &= 0, && \partial\Omega_N\times(0,T),\\ u &= 0, && \partial\Omega_D\times(0,T), \end{aligned}$$

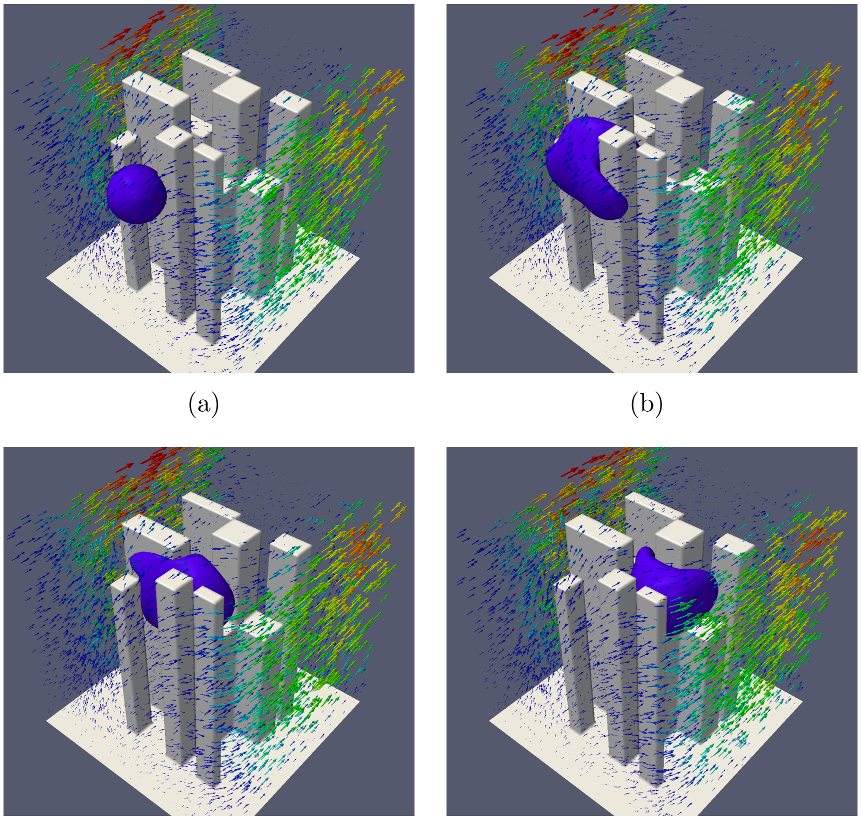
where u is the contaminant concentration, u_0 is the initial contaminant concentration field, $\boldsymbol{v}$ is the known wind velocity, k is the known diffusion coefficient, and T is the final time. Thus, the initial concentration u_0 represents the unknown parameter $\boldsymbol{x}$ and the sensor observations of concentration are the data $\boldsymbol{y}_{\text{obs}}$. Figure 5.1 describes the domain, boundary conditions, and velocity field for our example problem.

The code for statistical inversion extends an earlier code for the deterministic inverse problem [2, 1]. The discretization of the forward problem (5.1) uses a standard finite element approximation stabilized with SUPG (Streamline Upwind Petrov-Galerkin) to permit higher Peclet numbers. It uses linear tetrahedral elements and a uniform grid. Timestepping is implemented with Crank-Nicolson. Eigenvalues and eigenvectors are computed with Lanczos with periodic reorthogonalization [32]. PETSc[5, 6, 7] and SLEPc [33] are used throughout to enable parallelization.

As described in §2.1, the Bayesian formulation of a linear statistical inverse problem with Gaussian noise and prior is related to an appropriately-weighted least squares minimization problem. We choose to define our noise and prior pdf's by discretizing the infinite-dimensional functional

$$\frac{\beta_{\text{noise}}}{2}\int_{\Omega}\int_{0}^{T}(u-u_{\text{obs}})^2 b(x,t)\,d\mathbf{x}\,dt+\frac{\beta_{\text{prior}}}{2}\int_{\Omega}u_0^2\,d\mathbf{x}, \tag{5.2}$$

in which $u(x)$ satisfies the convection-diffusion equation (5.1). We study two alternatives for the observation operator $b(x,t)$. For the sparse observation case, measurements are taken at distinct points x_j, $j=0,\ldots,m-1$, over a



(a) (b)

(c) (d)

Figure 5.1: Time evolution of an atmospheric contaminant (deep blue) as it is transported through a model city composed of 10 buildings. The arrows represent wind velocity, and their size and color represent velocity magnitude. The mesh contains $112 \times 112 \times 112$ hexahedra, each of which is further subdivided into 6 tetrahedral elements. The velocity field was generated by solving the steady Navier-Stokes equations with a parabolic velocity inflow (from 0 to a maximum velocity of $v_{\max} = 1$), no-normal flow on the sides and top, traction-free outflow, and no-slip on the buildings and bottom. Additional parameters include density $\rho = 1$ and viscosity $\mu = 1$. The maximum velocity within the domain for the chosen parameters is 2.78. Contaminant boundary conditions are zero concentration on the inflow, ground, and buildings; and zero flux on the sides, outflow and top. The images depict snapshots of the contaminant concentration as well as flow field at the following times: (a) $T = 0$. (b) $T = 4$. (c) $T = 8$. (d) $T = 12$.

time interval $0 \leq t \leq T$. For the final time observation case, measurements are taken over the entire domain at final time $t = T$. The two observation operators are defined by

$$b(x,t) = \begin{cases} \sum_j \delta(x - x_j) & \text{sparse observations,} \\ \delta(t - T) & \text{final time observations.} \end{cases}$$

Discretization of expression (5.2) is equivalent to the choice of Gaussians pdf's for the prior and noise,

$$\pi_{\text{prior}}(\boldsymbol{u}_0) = \mathcal{N}(\bar{\boldsymbol{u}}_{0,\text{prior}}, \boldsymbol{\Gamma}_{\text{prior}}),$$

$$\pi_{\text{noise}}(\boldsymbol{e}) = \mathcal{N}(\bar{\boldsymbol{e}}, \boldsymbol{\Gamma}_{\text{noise}}),$$

with prior mean $\bar{\boldsymbol{u}}_{0,\text{prior}} = \boldsymbol{0}$ and noise mean $\bar{\boldsymbol{e}} = \boldsymbol{0}$, and prior and noise covariances given by

$$\boldsymbol{\Gamma}_{\text{prior}} = \frac{h^3}{\beta_{\text{prior}}}\boldsymbol{I},$$

$$\boldsymbol{\Gamma}_{\text{noise}} = \begin{cases} (h^3/\beta_{\text{noise}})\boldsymbol{I} & \text{final time observations,} \\ (1/\beta_{\text{noise}})\boldsymbol{I} & \text{sparse sensor observations,} \end{cases}$$

where β_{prior} and β_{noise} are weighting factors and h is the mesh size.

Discretization of the infinite-dimensional least squares minimization problem (5.2) results in the following expression

$$\frac{1}{2}(\boldsymbol{u} - \boldsymbol{u}_{\text{obs}})^T \boldsymbol{B}^T \boldsymbol{\Gamma}_{\text{noise}}^{-1} \boldsymbol{B}(\boldsymbol{u} - \boldsymbol{u}_{\text{obs}}) + \frac{1}{2}\boldsymbol{u}_0^T \boldsymbol{\Gamma}_{\text{prior}}^{-1} \boldsymbol{u}_0$$

where $\boldsymbol{B}$ is a discretization of $b(x,t)$, and the discretized contaminant field $\boldsymbol{u}$ satisfies $\boldsymbol{K}\boldsymbol{u} = \boldsymbol{C}\boldsymbol{u}_0$, which is the discretization of the forward convection-diffusion problem (5.1). Here, $\boldsymbol{K}$ is the discretized convection-diffusion operator, and $\boldsymbol{C}$ is the operator that maps initial conditions to space-time. Solving this equation for $\boldsymbol{u}$ given $\boldsymbol{u}_0$, we can rewrite the expression in terms of $\boldsymbol{u}_0$. Then the posterior pdf for the initial concentration field is given by the negative exponential of this least squares function, which yields the expression

$$\pi_{\text{post}}(\boldsymbol{u}_0) \propto \exp\left(-\frac{1}{2}(\boldsymbol{A}\boldsymbol{u}_0 - \boldsymbol{B}\boldsymbol{u}_{\text{obs}})^T \boldsymbol{\Gamma}_{\text{noise}}^{-1}(\boldsymbol{A}\boldsymbol{u}_0 - \boldsymbol{B}\boldsymbol{u}_{\text{obs}}) - \frac{1}{2}\boldsymbol{u}_0^T \boldsymbol{\Gamma}_{\text{prior}}^{-1}\boldsymbol{u}_0\right)$$

equivalent to (2.1) where $\boldsymbol{A} = \boldsymbol{B}\boldsymbol{K}^{-1}\boldsymbol{C}$ is the parameter-to-observable map. Since the problem is linear, the expression for the Hessian is

$$\boldsymbol{H} = \boldsymbol{A}^T \boldsymbol{\Gamma}_{\text{noise}}^{-1}\boldsymbol{A} + \boldsymbol{\Gamma}_{\text{prior}}^{-1} = \boldsymbol{H}_{\text{misfit}} + \boldsymbol{\Gamma}_{\text{prior}}^{-1}.$$

Forming the action of $\boldsymbol{A}$ on a given contaminant initial condition-like vector $\tilde{\boldsymbol{u}}_0$, as required by the low-rank approximation algorithm of §4.1, involves lifting the initial condition to space-time ($\boldsymbol{z} \leftarrow \boldsymbol{C}\tilde{\boldsymbol{u}}_0$), solution of the convection-diffusion problem for this source ($\boldsymbol{w} \leftarrow \boldsymbol{K}^{-1}\boldsymbol{z}$), and extraction of the resulting contaminant concentrations at sensor locations ($\boldsymbol{q} \leftarrow \boldsymbol{B}\boldsymbol{w}$). The action of the adjoint map $\boldsymbol{A}^T$ on the resulting weighted concentrations $\boldsymbol{\Gamma}_{\text{prior}}^{-1}\boldsymbol{q}$, as also required in the low rank approximation, involve the transpose of the operations described above and proceed in similar fashion.

5.2 Analysis of a model 1D convection-diffusion inverse problem

In this section we examine the structure of an analogous 1D infinite-dimensional convection-diffusion inverse problem to assess the effectiveness of a low-rank approximation. We describe the transport of a contaminant with the following convection-diffusion initial-boundary value problem:

$$\begin{aligned}
u_t - k\frac{\partial^2 u}{\partial x^2} + v\frac{\partial u}{\partial x} &= 0, && (0, L) \times (0, T), \\
u &= u_0, && (0, L) \times \{t = 0\}, \\
k\frac{\partial u}{\partial x}\bigg|_{x=0} &= k\frac{\partial u}{\partial x}\bigg|_{x=L}, && (0, T), \\
u|_{x=0} &= u|_{x=L}, && (0, T),
\end{aligned}$$

where the velocity v is taken as constant and the domain is periodic on the interval $(0, L)$. We assume the observation operator takes measurements at final time T of the concentration over the entire domain $(0, L)$. This model problem is a one-dimensional version of (5.1), but with periodic boundary conditions replacing the Dirichlet and Neumann conditions. The adjoint PDE problem is equivalent to the forward problem reversed in time with negative velocity, and is given by the terminal-boundary value problem

$$\begin{aligned}
-p_t - k\frac{\partial^2 p}{\partial x^2} - v\frac{\partial p}{\partial x} &= 0, && (0, L) \times (0, T), \\
p &= u(T), && (0, L) \times \{t = T\}, \\
k\frac{\partial p}{\partial x}\bigg|_{x=0} &= k\frac{\partial p}{\partial x}\bigg|_{x=L}, && (0, T), \\
p|_{x=0} &= p|_{x=L}, && (0, T).
\end{aligned}$$

We wish to investigate the viability of a low-rank approximation of the data misfit Hessian matrix corresponding to (5.2). For the finite-dimensional inverse problem, the prior-preconditioned data misfit Hessian matrix $\tilde{\boldsymbol{H}}_{\text{misfit}} = \boldsymbol{\Gamma}_{\text{prior}}^{1/2}\boldsymbol{A}^T\boldsymbol{\Gamma}_{\text{noise}}^{-1}\boldsymbol{A}\boldsymbol{\Gamma}_{\text{prior}}^{1/2}$ includes the influence of the prior and noise covariance matrices. As a surrogate for this, we will analyze, and verify the rapid decay in, the spectrum of the infinite-dimensional operator

$$\tilde{H}_{\text{misfit}} \overset{\text{def}}{=} \frac{\beta_{\text{noise}}}{\beta_{\text{prior}}} A^T A$$

where $A : u_0(x) \mapsto u(x,T)$ is the forward operator, $A^T : u(x,T) \mapsto p(x,0)$ is the adjoint operator, and $\beta_{\text{noise}}/\beta_{\text{prior}}$ represents the relative strength of the prior and the noise. Provided that the (finite dimensional) prior and noise covariance matrices do not corrupt the rapid decay in the spectrum seen in $\tilde{H}_{\text{misfit}}$, the low-rank approximation should still be appropriate. As mentioned before, a smoothing prior often reinforces this spectral decay. The expression for the eigenvalues of $\tilde{H}_{\text{misfit}}$ provides insight into the choice of prior and noise covariances that makes the low-rank approximation appropriate for a given statistical inverse problem. The choice of "spectrally neutral" i.i.d. prior and noise covariance matrices as in the previous subsection implies that the spectrum of $\tilde{\boldsymbol{H}}_{\text{misfit}}$ should still decay rapidly.

One can verify readily that the eigenfunctions of $\tilde{H}_{\text{misfit}}$ are given by

$$\cos\left(\frac{2\pi j x}{L}\right) \text{ and } \sin\left(\frac{2\pi j x}{L}\right) \text{ for } j = 1, 2, \ldots, \tag{5.3}$$

and the eigenvalues are given by

$$\frac{\beta_{\text{noise}}}{\beta_{\text{prior}}} e^{-8kT\pi^2 j^2/L^2} \text{ for } j = 1, 2, \ldots. \tag{5.4}$$

Note that the eigenvalues decay exponentially, which invites a low-rank approximation. The prior-preconditioned data misfit Hessian of the 3D inverse problem discussed in the next section is expected to exhibit similar structure, despite the effects of discretization, different boundary conditions, and complex geometry (as will be verified in the §5.3). Assuming the eigenvalues of $\tilde{\boldsymbol{H}}_{\text{misfit}}$ decay according to the expression (5.4), for a given cutoff $\lambda > \alpha$, the number of retained eigenvalues is therefore

$$r = 2 \left\lfloor \sqrt{\frac{L^2}{8kT\pi^2} \left| \log \left(\frac{\alpha \beta_{\text{prior}}}{\beta_{\text{noise}}} \right) \right|} \right\rfloor ,$$

where $0 < \alpha \leq \beta_{\text{noise}}/\beta_{\text{prior}}$. The subsections of §5.3 investigate thoroughly the properties of $\tilde{\boldsymbol{H}}_{\text{misfit}}$ for the 3D numerical example and support the analysis of the infinite-dimensional $\tilde{H}_{\text{misfit}}$ in this subsection.

We are also interested in the effect of physical parameters on the spectrum of $\tilde{H}_{\text{misfit}}$ and the corresponding eigenfunctions. As seen in (5.3), the eigenfunctions of the operator $\tilde{H}_{\text{misfit}}$ are oscillatory, where large eigenvalues are associated with smoother eigenfunctions and small eigenvalues are associated with increasingly oscillatory eigenfunctions. This is due to the effects of diffusion, which leads to greater loss of information in more oscillatory eigenfunctions (which are components of the initial concentration field), and therefore these eigenfunctions are harder to resolve from later-time observations. This behavior is explored for the 3D problem in §5.3.1.

Increasing the diffusion coefficient k leads to faster decay of the spectrum of $\tilde{H}_{\text{misfit}}$, as seen in (5.4). Moreover, a larger final time T also increases the rate of decay, since it allows more time for diffusion to act on components of the initial concentration. In this 1D example, there is no dependence of the spectrum of $\tilde{H}_{\text{misfit}}$ on the velocity v, since convection of the contaminant in the adjoint operator cancels (the oppositely-signed) convection in the forward operator due to periodicity. In the 3D numerical example, effects of convection may be visible, since with a more complex velocity field and non-periodic boundary conditions, convection does not cancel out between forward and adjoint PDEs. Effects of the physical parameters k and T, as well as additional experimental parameters, on the 3D numerical spectrum are further explored in §5.3.2.

We next analyze the effect of the choice of truncation of the spectrum in the low-rank approximation on the resulting error in the posterior covariance matrix. To facilitate the analysis, we assume the eigenvalues of the finite-dimensional $\tilde{\boldsymbol{H}}_{\text{misfit}}$ decay according to the infinite-dimensional expression (5.4). A bound on the squared error in the Frobenius norm (in both the approximation of $\tilde{\boldsymbol{H}}_{\text{misfit}}$ and the approximation of $\tilde{\boldsymbol{H}}_{\text{misfit}} + I$) in terms of r is derived by treating (4.6) as a right Riemann sums (the factor of two is from the repeated eigenvalues). Thus,

$$\begin{aligned} 2\sum_{j=r+1}^{n} \lambda_j^2 < 2\sum_{j=r+1}^{\infty} \lambda_j^2 &< \frac{2\beta_{\text{noise}}^2}{\beta_{\text{prior}}^2} \int_r^{\infty} e^{-16kT\pi^2 x^2/L^2} dx, \\ &< \frac{2\beta_{noise}^2}{\beta_{prior}^2} \sqrt{\frac{L^2}{32kT\pi}} \operatorname{erfc}\left(r\sqrt{\frac{16kT\pi^2}{L^2}} \right). \end{aligned}$$

Alternatively, a bound in terms of the cutoff α (such that retained eigenvalues satisfy $\lambda > \alpha$) can be derived. Let $R = \sqrt{\frac{L^2}{8kT\pi^2} \left| \log\left(\frac{\alpha\beta_{\text{prior}}}{\beta_{\text{noise}}} \right) \right|}$, so that $r = \lfloor R \rfloor$.

Then

$$\begin{aligned}
2\sum_{j=r+1}^{n}\lambda_j^2 &< \frac{2\beta_{\text{noise}}^2}{\beta_{\text{prior}}^2}\int_r^{\infty} e^{-16kT\pi^2x^2/L^2}dx,\\
&< 2\left(\lambda_R^2 + \frac{\beta_{\text{noise}}^2}{\beta_{\text{prior}}^2}\int_R^{\infty} e^{-16kT\pi^2x^2/L^2}dx\right),\\
&= 2\left(\lambda_R^2 + \frac{\beta_{\text{noise}}^2}{\beta_{\text{prior}}^2}\sqrt{\frac{L^2}{32kT\pi}}\text{erfc}\left(R\sqrt{\frac{16kT\pi^2}{L^2}}\right)\right),\\
&= 2\left(\alpha^2 + \frac{\beta_{\text{noise}}^2}{\beta_{\text{prior}}^2}\sqrt{\frac{L^2}{32kT\pi}}\text{erfc}\left(\sqrt{2\left|\log\left(\frac{\alpha\beta_{\text{prior}}}{\beta_{\text{noise}}}\right)\right|}\right)\right),\\
&< 2\left(1+\sqrt{\frac{L^2}{32kT\pi}}\right)\alpha^2,
\end{aligned}$$

where $0 < \alpha \leq \beta_{\text{noise}}/\beta_{\text{prior}}$. This expression is independent of the dimension of the parameter space, and therefore the number of retained eigenvalues in the low-rank approximation need not increase to maintain accuracy as the number of parameters increases.

5.3 Numerical experiments

We now study the behavior of the algorithm of §4.1 for the 3D convection-diffusion inverse problem described in §5.1. We examine the properties of the spectrum and eigenvectors of $\tilde{\boldsymbol{H}}_{\text{misfit}}$, and present experiments that demonstrate the effect of number of sensors, diffusion coefficient, and final time on the spectrum of $\tilde{\boldsymbol{H}}_{\text{misfit}}$. Finally, we examine the computational work associated with the low-rank approximation and demonstrate the scalability of the low-rank algorithm.

5.3.1 Properties of the spectrum of $\tilde{\boldsymbol{H}}_{\textbf{misfit}}$

In this section, we study the properties of eigenvalues and eigenvectors of $\tilde{\boldsymbol{H}}_{\text{misfit}}$.

Figure 5.2 presents the spectrum for the case of final time observations. The first 3500 eigenvalues of the spectrum of the $\tilde{\boldsymbol{H}}_{\text{misfit}}$ are shown in the figure. As can be seen, the eigenvalues rapidly collapse onto zero, and for this particular problem, 1500 modes (out of 274,625) are sufficient to capture the non-trivial structure of $\tilde{\boldsymbol{H}}_{\text{misfit}}$. As expressed by (4.5), the eigenpairs of the prior covariance matrix that are modified by observational data (i.e. the first 1500 or so) are precisely the dominant eigenpairs of $\tilde{\boldsymbol{H}}_{\text{misfit}}$ (i.e. the first 1500 or so eigenvalues in Figure 5.2). These are the directions for which the data (filtered through the prior) provide useful information. The remaining directions—eigenvectors 1500 through 274,625— for which the data (filtered through the prior) render no useful information need not be retained.

Next, we discuss the spectra for the case of sparse observations with 51 sensors and 274,625 uncertain parameters. In Figure 5.3 we present the spectrum and isosurfaces of selected eigenvectors of $\tilde{\boldsymbol{H}}_{\text{misfit}}$. Figure 5.3(a) depicts the largest 1800 eigenvalues of $\tilde{\boldsymbol{H}}_{\text{misfit}}$. Note that, like Figure 5.2, the eigenvalues decay rapidly and are negligible beyond the 500th (out of 274,625). However, the eigenvalue curve for this sparse observation case is discontinuous, exhibiting two visible jumps. The first jump occurs after the 51st eigenvalue, which is identical to the number of sensors. Indeed, each of the first 51 eigenvalues is associated with a unique sensor; this is illustrated by the 51st

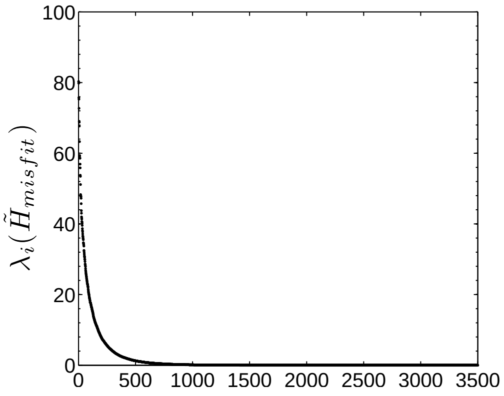


Figure 5.2: Spectrum of $\tilde{\boldsymbol{H}}_{\text{misfit}}$ for the final time observation case with final time $T = 6$, diffusion coefficient $k = 0.5$, $\beta_{\text{noise}}/\beta_{\text{prior}} = 100$, and a $64 \times 64 \times 64$ mesh. Largest 3500 (out of 274,625) eigenvalues of $\tilde{\boldsymbol{H}}_{\text{misfit}}$.

eigenvector shown in Figure 5.3(b), which depicts an isosurface in the shape of a ball surrounding the associated sensor. After the 51st eigenvalue, the associated eigenvectors change in shape by developing a large oscillation upstream of the sensor. The second jump in the spectrum occurs after the 101st eigenvalue, where the associated eigenvectors add a second oscillation upstream of the sensor. In §5.2, we noted that for a one-dimensional model problem with final-time observations, eigenvectors associated with smaller eigenvalues are increasingly oscillatory, reflecting the inability to reconstruct these modes due to diffusion-driven loss of information. This is also the case for time-dependent sparse observations within the complex geometry of the model city. However, in this case, the oscillations in the eigenvectors are not distributed evenly throughout the domain. They are focused on sensor locations, beginning with a ball surrounding a single sensor as in Figure 5.3(b), and becoming increasingly oscillatory and extending further upstream of the sensor, as seen in Figure 5.3(c)–5.3(e). Note that the oscillations in the eigenvectors are not uniform, likely due to the Neumann outflow boundary conditions and influence of convection on the prior-preconditioned Hessian of the data misfit.[1] These eigenvectors represent regions of influence for the sensors.

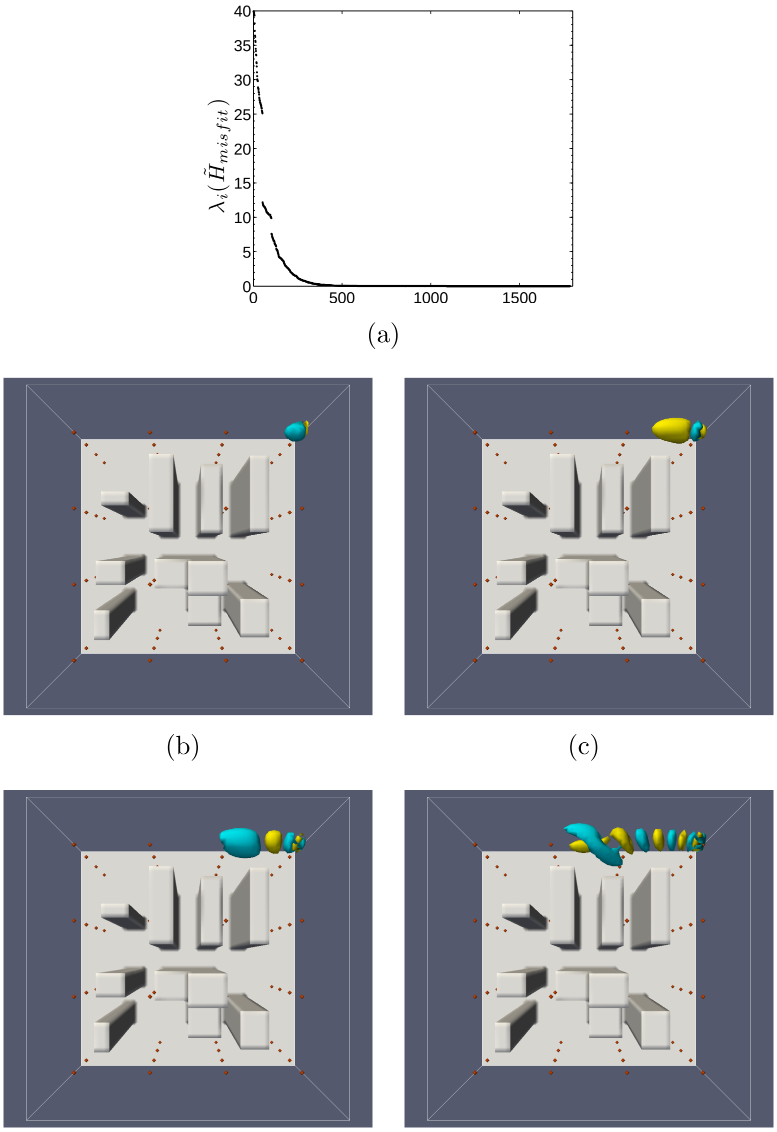
5.3.2 Dependence of spectrum of $\tilde{\boldsymbol{H}}_{\text{misfit}}$ on physical and experimental parameters

While we argued in §4.1 that refining the mesh for a class of ill-posed inverse problems does not affect the dominant eigenvalues/eigenvectors of the prior-preconditioned Hessian of the data misfit, other factors do affect the rate of decay of the spectrum and therefore the number of dominant eigenvalues. In this subsection, we study numerically the dependence of the spectrum of $\tilde{\boldsymbol{H}}_{\text{misfit}} = \boldsymbol{\Gamma}_{\text{prior}}^{1/2} \boldsymbol{A}^T \boldsymbol{\Gamma}_{\text{noise}}^{-1} \boldsymbol{A} \boldsymbol{\Gamma}_{\text{prior}}^{1/2}$ on number of sensors, diffusion coefficient, and final time.

In Figure 5.4(a), we examine the dependence of the spectrum on the number of sensors for the sparse sensor case. As can be seen in the figure, as the number of sensors increases, the eigenvalues decay less rapidly due to the information provided by the additional sensors. Thus, more eigenpairs should be retained to maintain accuracy of the approximation of $\tilde{\boldsymbol{H}}_{\text{misfit}}$. Table 5.1 presents the number of eigenvalues that need to be retained for a given accuracy, and the work required to capture them. This table demonstrates that

---

[1] Recall that in the analysis of the 1D model problem of §5.2 with periodic boundary conditions, convection in the forward and adjoint problems cancel each other, leading to the absence of convection in the data misfit Hessian.



(a)

(b)

(c)

(d)

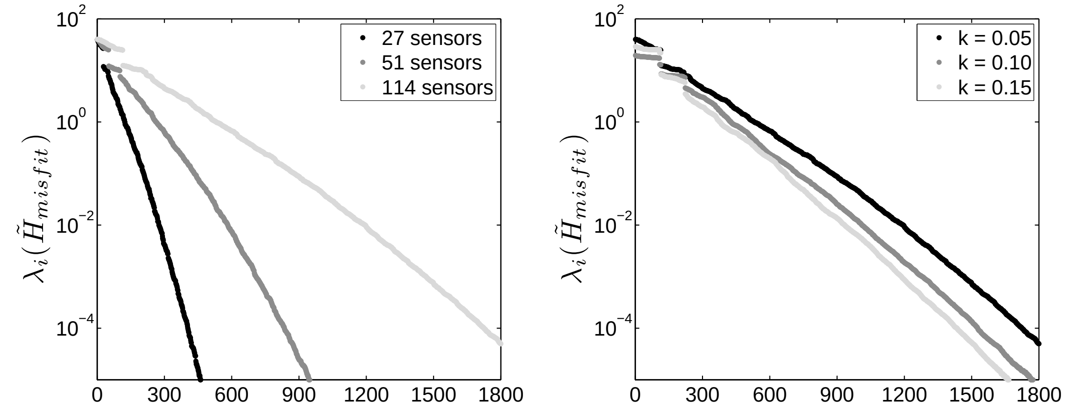
(e)

Figure 5.3: (a) Spectrum of $\tilde{\boldsymbol{H}}_{\text{misfit}}$ for a sparse sensor observation case with 51 sensors, final time $T = 8$, diffusion coefficient $k = 0.05$, $\beta_{\text{noise}}/\beta_{\text{prior}} = 100$, and a $64 \times 64 \times 64$ mesh. Largest 1800 (out of 274,625) eigenvalues of $\tilde{\boldsymbol{H}}_{\text{misfit}}$.are shown. (b)–(e) Isocontours of select eigenvectors of $\tilde{\boldsymbol{H}}_{\text{misfit}}$. Eigenvectors 51, 119, 185, and 359 are shown in (b), (c), (d), and (e), respectively. Red dots mark the locations of the sensors.

the low-rank approximation presented in §4.1 permits approximation of the Hessian for very large numbers of parameters at very low cost. For the case of 1,442,897 parameters, fewer than 0.1% of the eigenvalues are above the cutoff of 0.1, and thus an inexpensive approximation of the posterior parameter covariance can be readily made. For 27 sensors, 285 Hessian-vector products are required to capture the dominant portion of the spectrum, which as mentioned in §4.1 results in one forward/adjoint PDE solution per Hessian-vector product. This is dramatically fewer than the full 1.4 million forward/adjoint PDE solves needed to form the Hessian matrix, which would require years of supercomputing time. Note that while the number of dominant eigenvalues appears to increase linearly with the number of sensors, we cannot expect such a relationship in general. Here, the additional sensors are providing independent information due to the sparse placement of the sensors, while in general redundancy of sensor information may occur, permitting fewer retained eigenvalues and eigenvectors.

Figure 5.4(b) depicts the effect of changes in the diffusion coefficient on the spectrum, for a problem with 114 sensors. Increased diffusion results in less information captured by sensor observations, and so the spectrum decays more quickly. Thus, a lower rank approximation of $\tilde{\boldsymbol{H}}_{\text{misfit}}$ can be used.

The final time T also affects the spectrum as shown in Figure 5.5. For the sparse sensor case shown in Figure 5.5(a), measurements are taken at each timestep until the final time T, so that an increase in T leads to an increase in information provided by the sensors, thus requiring a larger number of retained



(a) (b)

Figure 5.4: Dependence of spectrum of $\tilde{\boldsymbol{H}}_{\text{misfit}}$ on sensor density and diffusion coefficient for the sparse observations case with final time $T = 8$, $\beta_{\text{noise}}/\beta_{\text{prior}} = 100$, and 274,625 initial concentration parameters. (a) Dependence of spectrum on the number of sensors, for diffusion coefficient $k = 0.05$. (b) Dependence of spectrum on the diffusion coefficient, for a problem with 114 sensors.

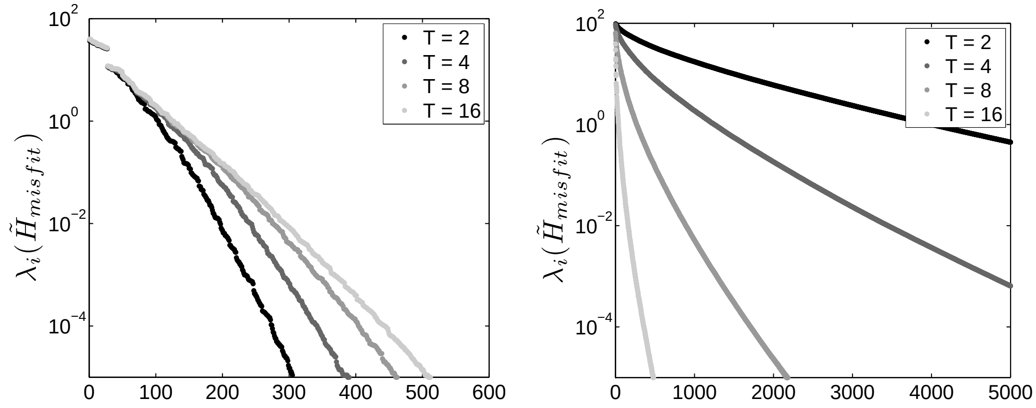
Number of sensors	Number of eigenvalues	Hessian-vector products
27	209	285
51	432	585
114	871	1130

Table 5.1: Influence of number of sensors on the work (in terms of Hessian-vector products) necessary to compute a low-rank approximation of $\tilde{\boldsymbol{H}}_{\text{misfit}}$ for an eigenvalue cutoff of $\lambda > 0.1$. Problem has 1,442,897 initial concentration parameters, diffusion coefficient $k = 0.05$, $\beta_{\text{noise}}/\beta_{\text{prior}} = 100$, and final time $T = 8$.

eigenpairs in the low-rank approximation. However, increasing the final time leads to diminishing returns, since as the initial contaminant field is transported out of the domain, fewer and fewer sensors provide information that can be used to infer the initial condition. Eventually, a longer time window $(0, T)$ will yield no new information, and the resulting curves in Figure 5.5(a) will tend toward each other. The diminishing returns are evident as we move from $T = 2$ to 4 to 8 to 16 in this figure.

For the final time observation case presented in Figure 5.5(b), increasing T has an opposite effect on the spectrum of $\tilde{\boldsymbol{H}}_{\text{misfit}}$, for two reasons. First, the longer we wait to observe, the more information is lost to diffusion, (which damps modes in relation to their wave number). Second, as T increases, more of the initial contaminant field is transported out of the range of the sensors, so that the amount of information provided by the sensors at final time T diminishes. These effects are evident in Figure 5.5(b). Note that this final time observation case most closely resembles the analytic results presented for the continuous 1D model problem in §5.2, although in that case the periodic boundary conditions prevented information loss due to convection.

Since a change in T affects all of the sensor data in the final time observation case, while it affects only some of the data in the sparse observation case, the influence of T on the spectrum is more pronounced in the former case, as can be seen by comparing Figures 5.5(a) and 5.5(b).



(a)

(b)

Figure 5.5: Effects of changing the final time T on the spectrum of $\tilde{\boldsymbol{H}}_{\text{misfit}}$ for the two sensor observation scenarios, for a problem with $k = 0.05$, $\beta_{\text{noise}}/\beta_{\text{prior}} = 100$, and 274,625 parameters. (a) Sparse sensor observation case (with 27 sensors). (b) Final time observation case.

5.3.3 Dependence of the spectrum of $\tilde{\boldsymbol{H}}_{\text{misfit}}$ on mesh size and resulting scalability of the low-rank algorithm

This section addresses the two properties required for scalability of the algorithm for low-rank-based approximation of the Hessian that was presented in §4.1. First, the number of eigenvalues of $\tilde{\boldsymbol{H}}_{\text{misfit}}$ that materially influence the Hessian should be independent of mesh size. Second, the work to compute those eigenpairs, measured in number of forward/adjoint solves of the 3D convection-diffusion PDE problems, should be independent of mesh size. Figure 5.6(a) presents the dependence of the spectrum of $\tilde{\boldsymbol{H}}_{\text{misfit}}$ on the mesh size, which is the same as the size of the parameter space. The figure shows the eigenvalue behavior for five successively finer mesh sizes, corresponding to a number of initial condition parameters ranging from 117,649 to 1,442,897.

As can be seen in the figure, increasing the size of the parameter space does not affect the spectrum of $\tilde{\boldsymbol{H}}_{\text{misfit}}$ substantially, and therefore does not change appreciably the number of eigenvalues required to approximate $\boldsymbol{\Gamma}_{\text{post}}$ up to a chosen eigenvalue cutoff. Indeed, a log-linear plot as in Figure 5.6(b) is needed to see the differences between the five spectra (note that the figure shows only the largest 0.07% to 0.9% of eigenvalues, depending on mesh size). In this figure, we see a small change in the spectrum as the mesh is refined for the smallest eigenvalues displayed in the plot, which as discussed in §5.3.1 are associated with more oscillatory eigenvectors. Eigenvectors associated with larger eigenvalues are smoother, and thus their corresponding eigenvalues are not affected by mesh refinement. On the other hand, the more oscillatory eigenvectors associated with smaller eigenvalues become better resolved with mesh refinement. The mesh should be fine enough to capture the eigenpairs that have an appreciable influence on the Hessian (and therefore should be retained in the low-rank approximation). The eigenvalue cutoff criterion provided in §4.1, namely that $\lambda \ll 1$, precisely achieves this goal. Figure 5.7(a), which displays the number of eigenvalues of $\tilde{\boldsymbol{H}}_{\text{misfit}}$ above a cutoff of $\lambda > 0.1$ as a function of number of parameters (and hence mesh size), demonstrates that the number of eigenvalues required to capture the dominant portion of the spectrum of $\tilde{\boldsymbol{H}}_{\text{misfit}}$ is independent of problem size, once the mesh is sufficiently fine to resolve the important (as identified from the data, filtered through the prior) eigenvectors.

The other property needed to insure scalability of our algorithm for

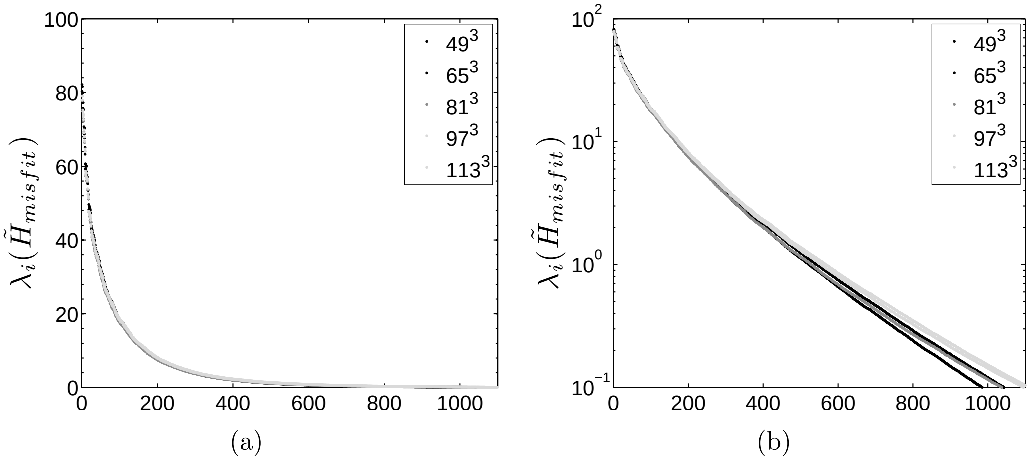
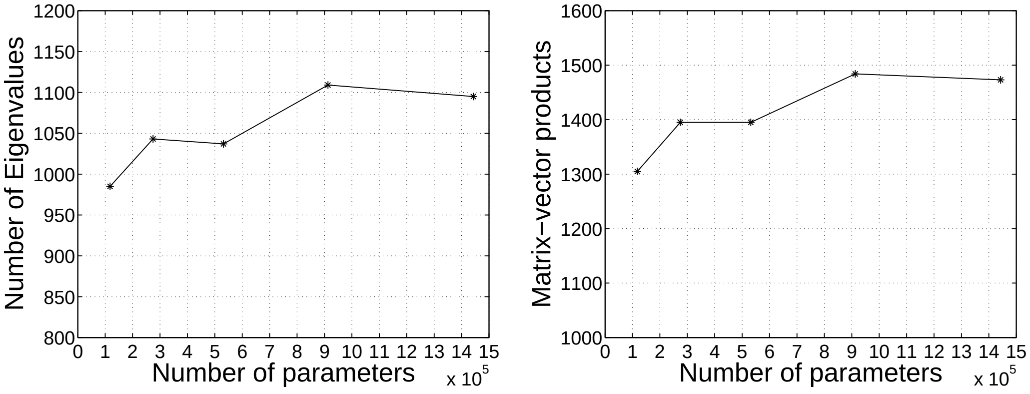


Figure 5.6: Dependence of the spectrum of $\tilde{\boldsymbol{H}}_{\text{misfit}}$ on mesh refinement for a final time observation case with $T = 6$, $\beta_{\text{noise}}/\beta_{\text{prior}} = 100$, and diffusion coefficient $k = 0.05$. (a) Eigenvalues of $\tilde{\boldsymbol{H}}_{\text{misfit}}$ such that $\lambda > 0.1$ for different mesh resolution. (b) Log-linear plot of eigenvalues of $\tilde{\boldsymbol{H}}_{\text{misfit}}$ such that $\lambda > 0.1$ for different mesh resolution.

low-rank estimation of the posterior covariance is that the work required to capture the dominant eigenvalues also does not depend on mesh size. Towards this end, we choose to extract the dominant eigenvalues of $\tilde{\boldsymbol{H}}_{\text{misfit}}$ using the Lanczos algorithm option in SLEPc. The main computational kernel in the Lanczos algorithm is to form a product of the matrix $\tilde{\boldsymbol{H}}_{\text{misfit}}$ with a vector. As discussed in §4.1, $\tilde{\boldsymbol{H}}_{\text{misfit}}$ is never explicitly formed; instead, each matrix-vector product is dominated by one forward and one adjoint solve of the convection-diffusion PDE. The number of matrix-vector products is thus a good measure of the cost of the low-rank approximation: solution of the forward and adjoint PDEs overwhelms all other components, including in the eigenvalue computation. Figure 5.7(b) depicts the number of matrix-vector products (and therefore forward/adjoint PDE solves) to compute the retained eigenvalues shown in Figure 5.7(a). The parameters again range in number from 117,649 to 1,442,897. As can be seen, the number of required matrix-vector products is again largely independent of parameter dimension. In summary, Figure 5.7(b) illustrates that the Hessian can be estimated accurately at a cost (measured in PDE solves) that is independent of problem size.



(a) (b)

Figure 5.7: Work required to estimate dominant spectrum of $\tilde{\boldsymbol{H}}_{\text{misfit}}$ as a function of mesh size/number of parameters, for a final time observation case with $T = 6$, $\beta_{\text{noise}}/\beta_{\text{prior}} = 100$, and diffusion coefficient $k = 0.05$. (a) Number of eigenvalues of $\tilde{\boldsymbol{H}}_{\text{misfit}}$ retained, such that $\lambda > 0.1$. (b) Number of matrix-vector products (and hence forward/adjoint PDE solves) required to compute eigenvalues of $\tilde{\boldsymbol{H}}_{\text{misfit}}$ such that $\lambda > 0.1$.

5.4 Spectrum of the posterior covariance

We continue from the expression for the posterior pdf (2.1) for a nonlinear inverse problem in §2.1. Let us now assume that the parameter-to-observable map is linear, i.e.

$$f(X) = \boldsymbol{A}X.$$

Here, $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ is the linear operator that maps parameters $\boldsymbol{x}$ to observables $\boldsymbol{y}$, via the solution of a large-scale discretized PDE problem. In this case, $\pi_{\text{post}}(\boldsymbol{x})$ is also Gaussian, with mean $\bar{\boldsymbol{x}}_{\text{post}} \in \mathbb{R}^n$ given by the maximum *a posteriori* (MAP) point, i.e.,

$$\bar{\boldsymbol{x}}_{\text{post}} = \arg\max_{\boldsymbol{x}} \pi_{\text{post}}(\boldsymbol{x}).$$

Finding the MAP point is equivalent to solving a weighted least squares optimization problem, i.e.,

$$\bar{\boldsymbol{x}}_{\text{post}} = \arg\min_{\boldsymbol{x}} \left(\frac{1}{2} \|\boldsymbol{y}_{\text{obs}} - \boldsymbol{A}\boldsymbol{x} - \bar{\boldsymbol{e}}\|^2_{\boldsymbol{\Gamma}^{-1}_{\text{noise}}} + \frac{1}{2} \|\boldsymbol{x} - \bar{\boldsymbol{x}}_{\text{prior}}\|^2_{\boldsymbol{\Gamma}^{-1}_{\text{prior}}} \right), \tag{5.5}$$

which amounts to solving a regularized deterministic inverse problem, where $\boldsymbol{\Gamma}^{-1}_{\text{prior}}$ plays the role of the regularization operator, and $\boldsymbol{\Gamma}^{-1}_{\text{noise}}$ is a weighting of the data misfit term. The covariance matrix of the posterior pdf of model parameters, $\boldsymbol{\Gamma}_{\text{post}} \in \mathbb{R}^{n\times n}$, is given simply by the inverse of the Hessian matrix of the least squares objective function, i.e.,

$$\boldsymbol{\Gamma}_{\text{post}} = \left(\boldsymbol{A}^T \boldsymbol{\Gamma}^{-1}_{\text{noise}} \boldsymbol{A} + \boldsymbol{\Gamma}^{-1}_{\text{prior}} \right)^{-1}. \tag{5.6}$$

So, in summary,

$$\pi_{\text{post}}(\boldsymbol{x}) = \mathcal{N}(\bar{\boldsymbol{x}}_{\text{post}}, \boldsymbol{\Gamma}_{\text{post}}).$$

Here we examine properties of the posterior covariance matrix rather than focusing on scaling of the low-rank approximation. In addition, by (4.11), the analysis of the error bound for low-rank truncation error for a 1D periodic model problem in §5.2 of an infinite dimensional operator still applies. The following sections will examine the spectrum of the posterior covariance matrix, the variance of the posterior pdf, and the effect of the low-rank approximation on the accuracy of the variance.

In this section, we study the properties of the eigenvalues of the approximate posterior covariance matrix.

Figure 5.8(a) presents the spectrum for the case of final time observations. The first 3500 eigenvalues of the posterior covariance matrix (based on a 3500-eigenpair low-rank approximation of $\tilde{\boldsymbol{H}}_{\text{misfit}}$) are displayed in Figure 5.8(a). The asymptotic horizontal line of the curve reflects prior information, which in this case is a scaled standard normal pdf. The information provided by observational data is responsible for the departure from the horizontal line seen in this figure. As expressed by (4.10), the eigenpairs of the prior covariance matrix that are modified by observational data (i.e. the first 1500 or so) are precisely the dominant eigenpairs of $\tilde{\boldsymbol{H}}_{\text{misfit}}$ (i.e. the first 1500 or so eigenvalues in Figure 5.2). The posterior pdf thus inherits its uncertainty from the prior in those directions for which the data (filtered through the prior) render no useful information (these are eigenvectors 1500 through 274,625). The remaining directions—for which the data (filtered through the prior) do provide useful information—serve to reduce the uncertainty in the posterior.

Next, we discuss the spectra for the case of sparse observations with

51 sensors and 274,625 uncertain parameters. In Figure 5.3 we presented the spectrum and isosurfaces of selected eigenvectors of $\tilde{\boldsymbol{H}}_{\text{misfit}}$. Figure 5.3(a) depicts the largest 1800 eigenvalues of $\tilde{\boldsymbol{H}}_{\text{misfit}}$, while Figure 5.8(b) shows the corresponding spectrum of the posterior covariance matrix based on this low-rank approximation. Similar to the final time observation case, information from observational data lowers the eigenvalues of the sparse-observation posterior covariance matrix associated with dominant eigenpairs of $\tilde{\boldsymbol{H}}_{\text{misfit}}$, thereby reducing the first 500 or so eigenvalues in Figure 5.8(b). However, the reduction in uncertainty (i.e. the departure from the horizontal line in the figure) is largest around the sensors themselves, as illustrated by the eigenvector depicted in Figure 5.3(b). These eigenvectors represent regions of influence for the sensors; these regions affect how the prior covariance matrix is modified in (4.10) to yield the posterior covariance estimate.

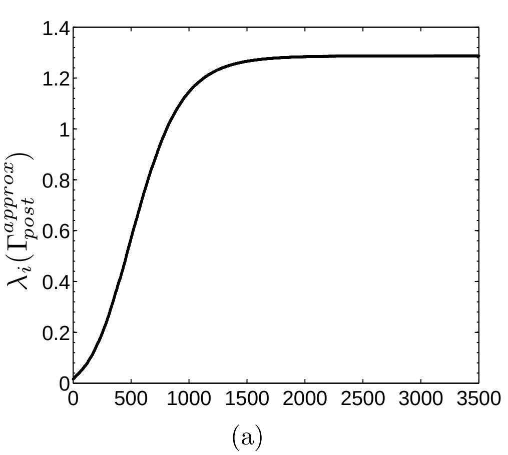
5.5 Interpretation of the posterior variance

The approximation of the posterior covariance is given by (4.10). In this section, we study the approximated variance. For the present example, the specific choice of $\boldsymbol{\Gamma}_{\text{prior}}$ and $\boldsymbol{\Gamma}_{\text{noise}}$ are given in §5.1 as scaled identities, leading to

$$\boldsymbol{\Gamma}_{\text{post}} \approx \frac{h^3}{\beta_{\text{prior}}}(\boldsymbol{I} - \boldsymbol{V}_r \boldsymbol{D}_r \boldsymbol{V}_r^T).$$

Recall that $\boldsymbol{V}_r$ and $\boldsymbol{D}_r$ represent the dominant eigenpairs of the prior-preconditioned Hessian matrix of the data misfit, $\tilde{\boldsymbol{H}}_{\text{misfit}}$.

In Figure 5.9, we plot the approximate posterior variance of the initial

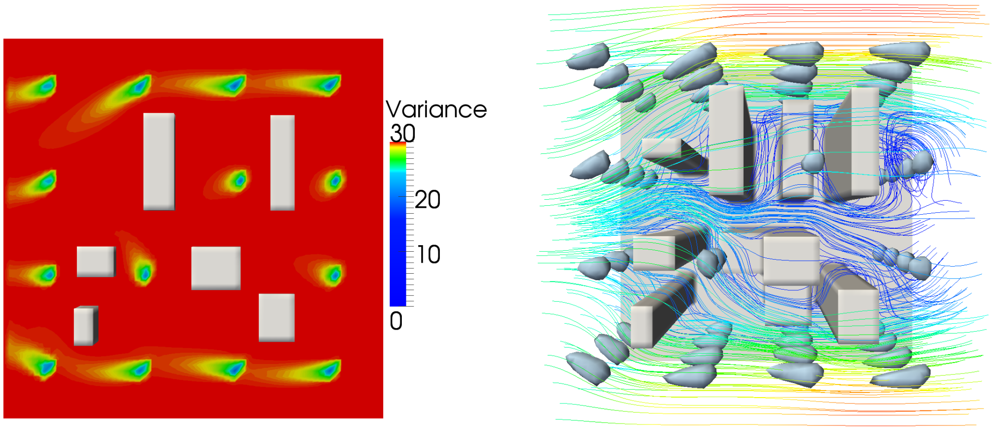


(a)

(b)

Figure 5.8: (a) Spectrum of the posterior covariance estimate for the final time observation case with final time $T = 6$, diffusion coefficient $k = 0.5$, $\beta_{\text{noise}}/\beta_{\text{prior}} = 100$, and a $64 \times 64 \times 64$ mesh. (a) 3500 (out of 274,625) eigenvalues of the approximation of the posterior covariance $\mathbf{\Gamma}_{\text{post}}$ corresponding to the largest 3500 (out of 274,625) eigenvalues of $\tilde{\boldsymbol{H}}_{\text{misfit}}$.
(b) Spectrum of the approximated posterior covariance for a sparse sensor observation case with 51 sensors, final time $T = 8$, diffusion coefficient $k = 0.05$, $\beta_{\text{noise}}/\beta_{\text{prior}} = 100$, and a $64 \times 64 \times 64$ mesh. Corresponding 1800 eigenvalues of $\mathbf{\Gamma}_{\text{post}}$ to the largest 1800 (out of 274,625) eigenvalues of $\tilde{\boldsymbol{H}}_{\text{misfit}}$ are shown.

concentration field, which is given by the diagonal of $\mathbf{\Gamma}_{\text{post}}$, for the case of 51 sensors distributed uniformly (avoiding the building interiors) throughout the model city. Figure 5.9(a) presents a cross-section of the variance field, taken horizontally through the top row of sensors. Recall that the prior estimate of the variance is h^3/β_{prior}, and the posterior estimate modifies this value according to the expression (4.10). The bright red area is the region of the domain where the prior dominates. Information gained from observations does not reduce the uncertainty in the contaminant concentration in these regions. The influence of the observations on the posterior variance is concentrated around the sensors, as expected. The lowest uncertainty (dark blue) is exactly at the sensors, with increasing uncertainty moving away from the sensor locations. A three-dimensional view of the variance is given in Figure 5.9(b), which shows isocontours of the variance in light blue in the context of the surrounding buildings, superposed on streamlines of the velocity field. The isocontours capture the low uncertainty region surrounding each sensor. Note that these low variance regions vary in shape. In the leeward regions of buildings, where diffusion dominates, the low variance regions are more isotropic. Where convection dominates, the low variance regions associated with each sensor extend anisotropically upstream, following the streamlines of the velocity field. Contaminant that is initially directly upstream of a sensor is convected along streamlines and measured by downstream sensors, thereby reducing uncertainty of the reconstructed initial concentration in these upstream regions. On the other hand, contaminants downstream of a sensor tend to experience greater diffusion as they are convected a further distance before encountering a downstream sensor. Thus, the stronger the convection, the longer the low variance tails. In low-velocity regions where contaminants spread mainly through diffusion, however, the variance reduces in a more isotropic fashion. The alignment of the anisotropy of the variance with streamlines is evident in Figure 5.9.

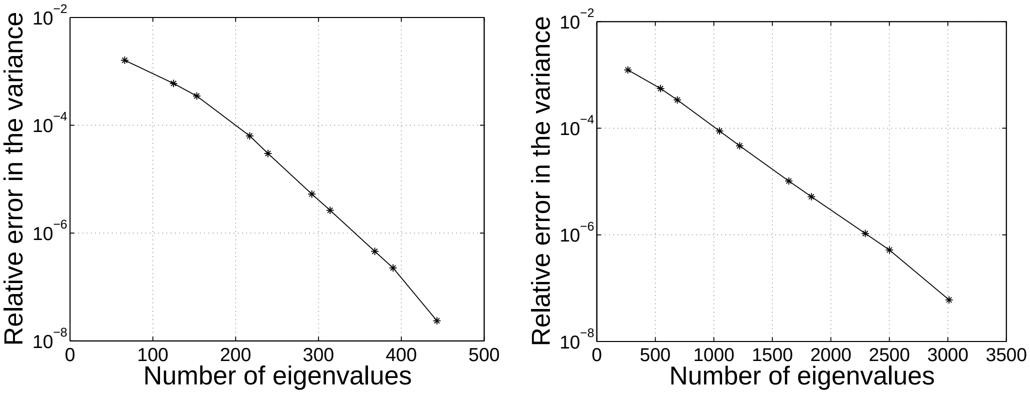


(a) (b)

Figure 5.9: The variance in the initial concentration of the contaminant, for the sparse observation case of 51 sensors measuring to final time $T = 8$ with diffusion coefficient $k = 0.03$ for a $64 \times 64 \times 64$ mesh. The covariance approximation is based on an eigenvalue cutoff of $\lambda > 0.1$, and required 513 eigenvalues out of 274,625, computed at a cost of 650 forward/adjoint convection-diffusion PDE solutions. We chose $\beta_{\text{prior}} = h^3/30$ and $\beta_{\text{noise}} = 100\beta_{\text{prior}}$. (a) Cross-section of the variance field through the top row of sensors, with arrows representing the velocity field. (b) Top view, isocontours of the variance field superposed on streamlines of the velocity field that are colored by magnitude.

5.6 Effect of low-rank approximation of $\tilde{\boldsymbol{H}}_{\text{misfit}}$ on accuracy of variance approximation

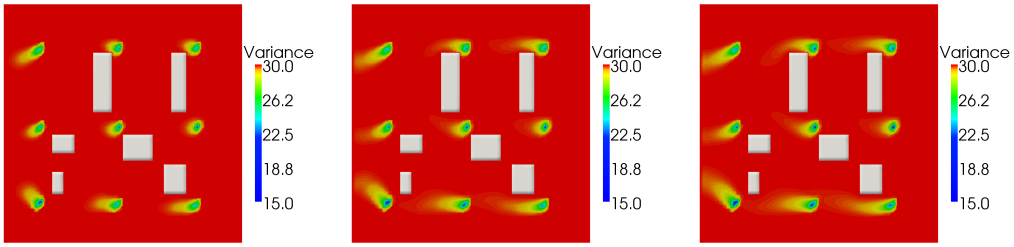
In this section, we study the effect of truncation of the spectrum, and resulting low-rank approximation of $\tilde{\boldsymbol{H}}_{\text{misfit}}$, on the solution of the statistical inverse problem, specifically the evaluation of the main diagonal of the posterior covariance matrix, i.e. the variances of the initial concentration parameters. Figure 5.10 presents the relative error in the approximated variance as a function of the number of eigenvectors retained. As can be seen in the figure, the relative error decreases rapidly with the number of retained eigenvectors of $\tilde{\boldsymbol{H}}_{\text{misfit}}$. For both the sparse observation case and the final time observation case, inclusion of additional eigenpairs in the approximation of Γ_{post} has diminishing returns with respect to increased accuracy.



(a) (b)

Figure 5.10: Dependence of relative error in the approximated variance on the number of eigenvalues retained in the approximation. Since the dimension of $\tilde{\boldsymbol{H}}_{\text{misfit}}$ is 274,625, computing the exact variance at each grid point, $\text{Var}_i^{\text{exact}}$, is prohibitive; instead, the "exact" solution is defined using an aggressive eigenvalue cutoff of $\lambda > 10^{-5}$. The relative error is defined by $\sum_i |\text{Var}_i^{\text{exact}} - \text{Var}_i^{\text{approx}}| / \sum_i |\text{Var}_i^{\text{exact}}|$. (a) Relative error in the variance for a sparse observation case with 27 sensors, final time $T = 16$, $\beta_{\text{noise}}/\beta_{\text{prior}} = 100$, and diffusion coefficient $k = 0.05$. The "exact" variance retains 462 eigenpairs in the approximation of $\tilde{\boldsymbol{H}}_{\text{misfit}}$. (b) Relative error in the variance for a final time observation case with final time $T = 6$ and diffusion coefficient $k = 0.05$. The "exact" variance retains 3,240 eigenpairs.

These diminishing returns are also illustrated in Figure 5.11, which displays the effects of truncation on the approximate variance field (on a horizontal plane) for a sparse observation case with 27 sensors. This is the same model city as the one used to generate Figure 5.10(a). The sequence of three images in the figure show the improvement in the approximation of the variance field as the number of retained eigenvectors is increased corresponding to a eigenvalue cutoff of 5, 1, and 0.1. The differences are most clearly visible in the increased length of the low variance tails (yellow/green) upstream of the sensors. These are the regions for which the observations still provide some information (when filtered through the prior), before yielding to the red areas, which are determined only by the prior. The lowest variance areas (in blue, at the sensor locations) extend further for the smallest cutoff value. While the 59 additional eigenvectors used in the variance approximation of Figure 5.11(b) compared to Figure 5.11(a) provide visible improvement, the additional 92 eigenvectors needed to generate the approximate variance field in Figure 5.11(c) provide negligible returns in accuracy.



(a) (b) (c)

Figure 5.11: The variance in the initial concentration of the contaminant, based on three approximations to Γ_{post} from different truncation cutoff values for a sparse observation case with 27 sensors, final time $T = 16$, diffusion coefficient $k = 0.05$, and 274,625 parameters. (a) Cutoff $\lambda > 5$: 66 eigenvalues. (b) Cutoff $\lambda > 1$: 125 eigenvalues. (c) Cutoff $\lambda > 0.1$: 217 eigenvalues.

Chapter 6

A 2D groundwater inverse problem

We now turn to the solution of the statistical inverse problem: we want to calculate properties of the posterior pdf, such as the mean, variance, and marginals. We will do this through the MCMC algorithms presented in §2.3.

6.1 Inverse problem to describe the log transmissivity as characterized by 121 parameters

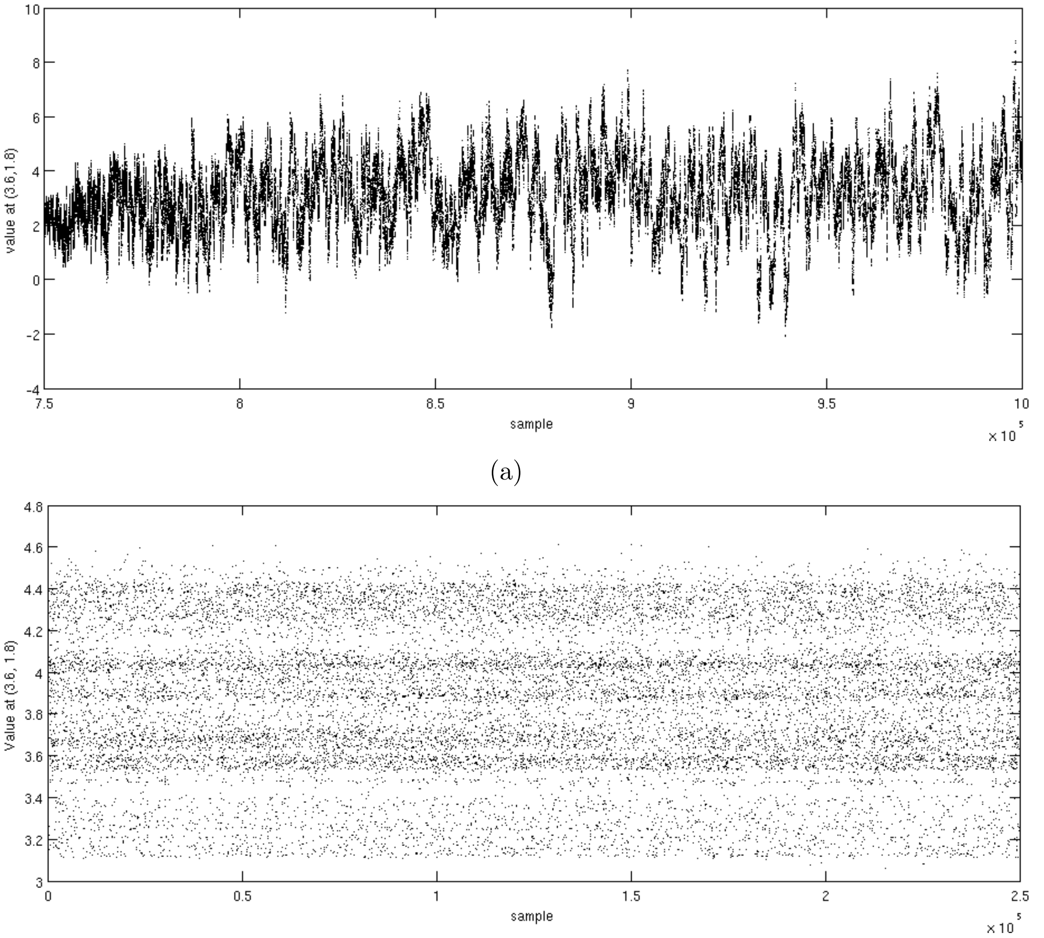
In this section, we use the algorithms described in Chapter 2 and Chapter 4 to solve the groundwater flow inverse problem initially presented in §4.3.2. We begin with a presentation of some of the properties of the posterior pdf in §6.1.1, then turn to MCMC diagnostics and the dependence of the algorithm on the number of greedy iterations in §6.1.2. In §6.1.3 we compare these results with those from sampling the response surface as a surrogate for the posterior. §6.1.4 continues with results based on the two-stage delayed acceptance algorithm. Finally, §6.1.5 compares the costs and benefits of the three algorithms.

6.1.1 Properties of the posterior pdf

We first introduce the solution to the deterministic inverse problem – the MAP point – in Figure 6.2(b). Despite relying on only 9 noisy observations of the head (see description of synthetic data in §4.3.2), we have recovered the channel feature in the center of the domain in Figure 6.2(a). Other features, such as the channel to the left and the low transmissivity region to the right, do not appear in this estimate of the log transmissivity.

As discussed before, we are seeking a solution to the statistical inverse problem, which provides not only an estimate of the log transmissivity, but a full posterior pdf describing the uncertainty in the parameter field. In this section, we will present properties of this solution based on Metropolis-Hastings sampling of the posterior pdf with the Gaussian mixture proposal generated by 75 greedy iterations. We attempted to provide a comparison set of results through the Delayed Rejection Adaptive Metropolis (DRAM) algorithm [30], however, this method has difficulty with high-dimensional parameter spaces (as all contemporary MCMC methods do) and failed to converge. In Figure 6.1(a), we present a traceplot from the end of a 1,000,000 sample chain, which shows that the samples are still highly correlated and the chain is not well-mixed. Therefore, we will provide various diagnostics indicating the strength of our own MCMC algorithms. For instance, the traceplot from sampling the posterior pdf with the Gaussian mixture proposal is shown in Figure 6.1(b), and displays the characteristic fuzzy appearance of a well-mixed chain.

Thus, in Figure 6.2(c) we present the mean of the posterior pdf cal-



(a)

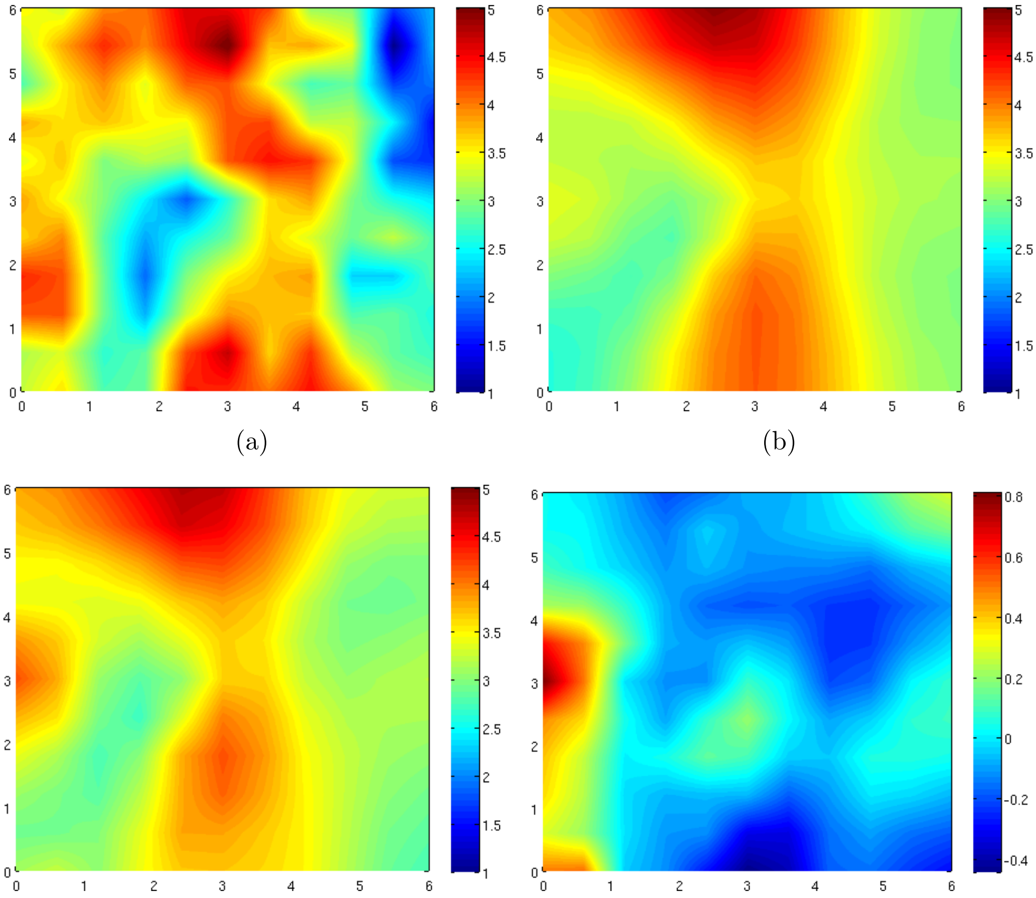
(b)

Figure 6.1: (a) Traceplot of a chain generated through DRAM. The value of the 40th degree of freedom (corresponding to the point $(3.6, 1.8)$ for the last 250,000 samples in a 1,000,000 sample chain starting with the MAP point, given the inverse of the local Hessian approximation as an estimate of the covariance. (b) Traceplot of a 250,000 sample chain starting with the MAP point. The chain was generated through Metropolis-Hastings sampling of the posterior pdf with the Gaussian mixture proposal generated by 75 greedy iterations.

culated through Metropolis-Hastings sampling of the posterior pdf with the Gaussian mixture proposal generated by 75 greedy iterations. Figure 6.2(c) shows the differences between the mean and the MAP. Note that the mean differs from the MAP point (particularly at the left of the domain, but also in the width of the channel). The posterior pdf is therefore non-Gaussian. In fact, this mean is a slightly better estimate of the synthetic field: the L^2 error from the mean is 3.2027 and from the MAP is 3.3051.

We next consider the variance of the posterior pdf (Figure 6.3(c)). For comparison, the variance of the prior pdf can be seen in Figure 6.3(a), and the variance based on a local Gaussian approximation to the posterior pdf at the MAP point in Figure 6.3(b). We see that the additional information from the observations and the forward model results in a much lower and more irregular variance for the posterior pdf relative to the prior pdf. The variance is highest near $(0, 0)$, and lowest in the region around and between the 9 observation points. This figure also emphasizes that a Gaussian approximation at the MAP point is a poor approximation to the posterior pdf. To further explore the implications of the variance, Figure 6.1.1 shows the mean plus and minus one standard deviation.

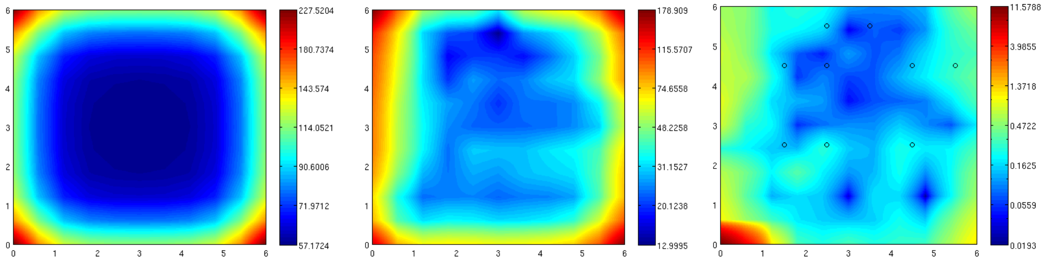
We turn to Figure 6.5 and Figure 6.6, for additional intuition on the differences between the prior pdf and the posterior pdf. In Figure 6.5 we see samples drawn from the prior pdf, and the associated head determined by the forward model. Our prior information (as described by the prior pdf) suggests the log transmissivity field contains a random scattering of lumps and dips



(a) (b)

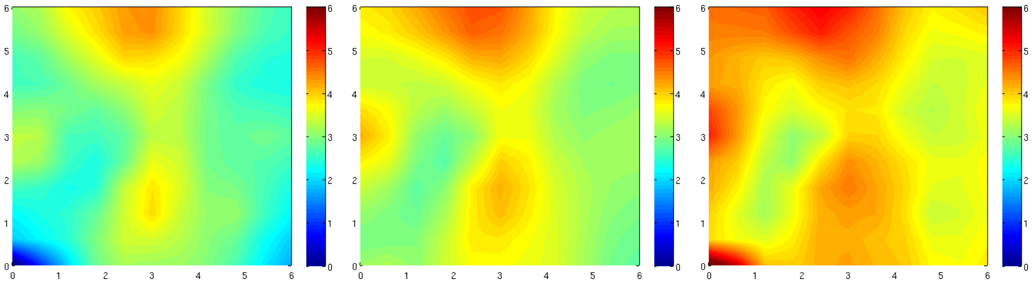
(c) (d)

Figure 6.2: (a) Synthetic log transmissivity field used to generate observations (see §4.3.2). (b) Solution of the deterministic inverse problem: the MAP point (log transmissivity field). (c) Mean of posterior pdf as calculated by Metropolis-Hastings sampling of the posterior pdf with the Gaussian mixture proposal generated by 75 greedy iterations. (d) Mean of posterior pdf minus the MAP point.



(a) (b) (c)

Figure 6.3: (a) Variance of the prior parameter pdf. (b) Approximate variance based on a local Gaussian approximation at the MAP point. (c) Variance of posterior pdf as calculated by Metropolis-Hastings sampling of the posterior pdf with the Gaussian mixture proposal generated by 75 greedy iterations. Observation points marked by black circles.



(a) (b) (c)

Figure 6.4: Results calculated by Metropolis-Hastings sampling of the posterior pdf with the Gaussian mixture proposal generated by 75 greedy iterations. (a) One standard deviation below the mean. (b) Mean. (c) One standard deviation above the mean.

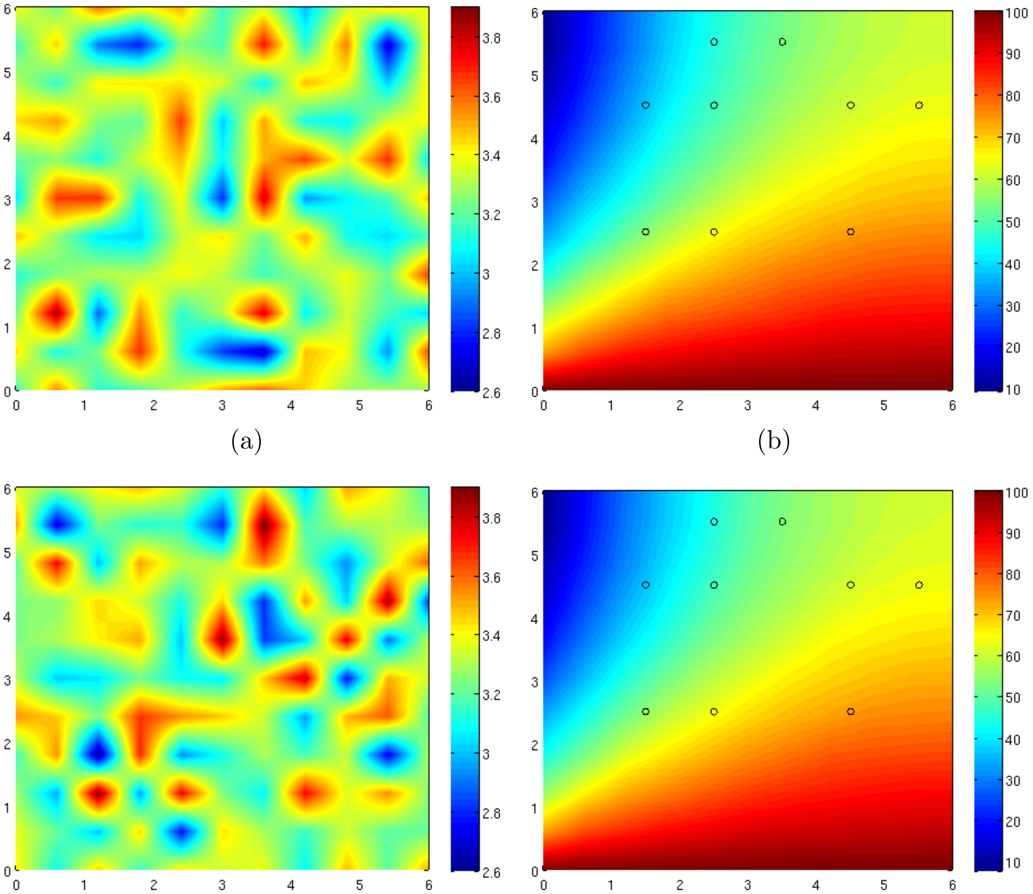
centered around the prior mean value (which is a constant). In contrast, the samples of the posterior pdf in Figure 6.6, show a distinct structure. The posterior pdf combines the prior pdf with information from observations and the forward model. Indeed, many of these samples show a highly permeable channel down the center (which corresponds to information derived from the synthetic field observations and the likelihood).

Finally, we present Figure 6.7, which shows two additional shape points (the MAP point is also a shape point). Their influence can be seen in the samples in Figure 6.6. It also shows two cross-sections, where we can see the actual values of the posterior pdf compared to the Gaussian mixture and Gaussian process response surface.

6.1.2 Metropolis-Hastings sampling of the posterior pdf with the Gaussian mixture proposal

In this section, we describe in more detail the dependence of the results in the previous section on the number of greedy iterations used to build the Gaussian mixture proposal, and also provide some diagnostic measures indicating convergence of the MCMC chain. We begin with the traceplot in Figure 6.8 based on the value of the samples in the chain at $(3.6, 1.8)$. The striped nature of the plot is due to the multimodal nature of the pdf. The chain looks well-mixed, and the samples do not look correlated.

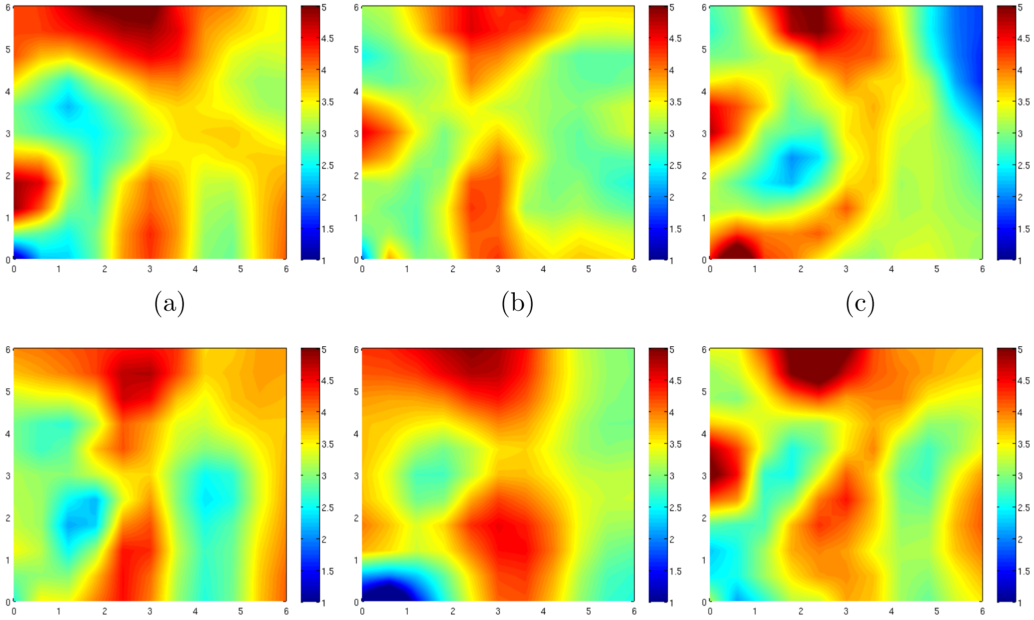
In order to investigate convergence of the algorithms, for each algorithm we produce an initial chain starting at the MAP point of 250,000 samples.



(a) (b)

(c) (d)

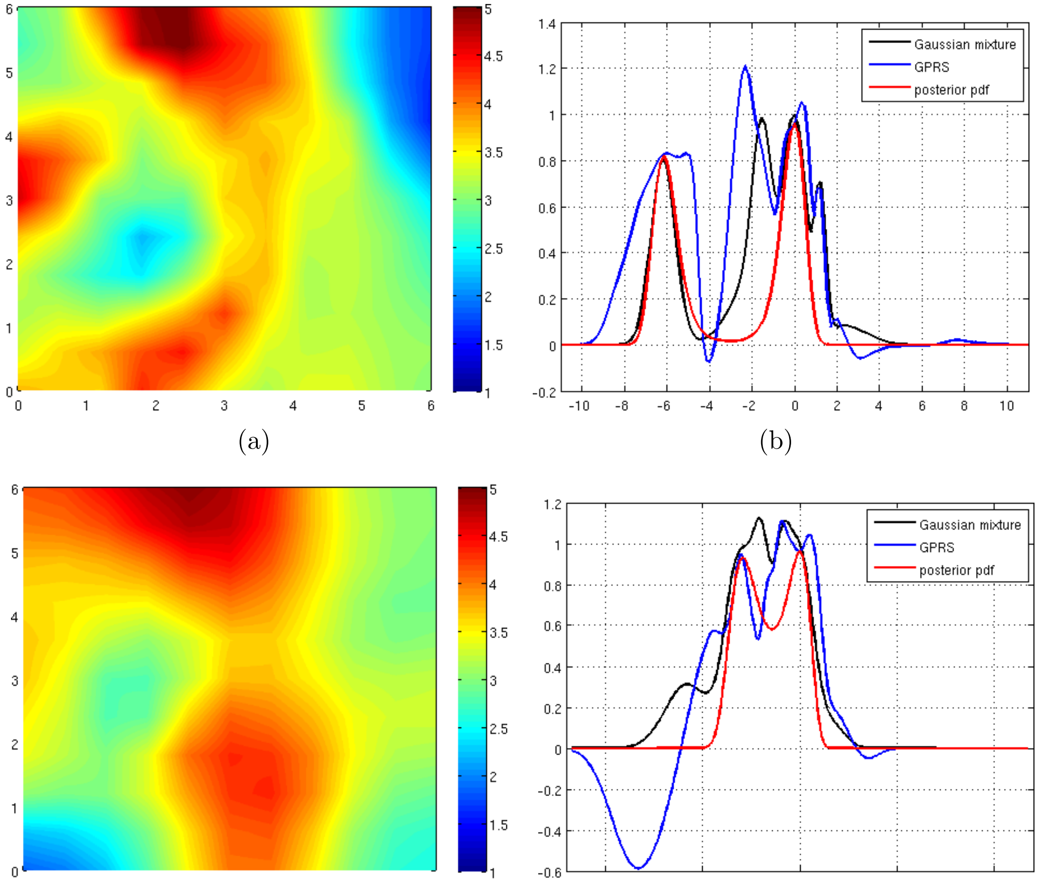
Figure 6.5: The prior parameter pdf describes the properties of the log transmissivity field by a Gaussian with a constant mean 3.290624 (the average value of the synthetic log transmissivity field in §4.3.2) and a prior covariance as described in Chapter 3 with parameters 0.05 and 0.03. (a) Sample drawn from the prior parameter pdf. (b) Head based on (a) and forward model. (c) Sample drawn from the prior parameter pdf. (d) Head based on (c) and forward model. Black circles mark observation points.



(a) (b) (c)

(d) (e) (f)

Figure 6.6: A selection of samples of the posterior pdf, generated through Metropolis-Hastings sampling of the posterior pdf with the Gaussian mixture proposal generated by 75 greedy iterations.



(a) (b)

(c) (d)

Figure 6.7: (a) Second shape point. (b) Cross-section of the posterior pdf, GPRS, and Gaussian mixture after 25 greedy iterations, along the line passing through the MAP point (the first shape point) and the second shape point. (c) Third shape point. (d) Cross-section of the posterior pdf, GPRS, and Gaussian mixture after 25 greedy iterations, along the line passing through the MAP point (the first shape point) and the third shape point.

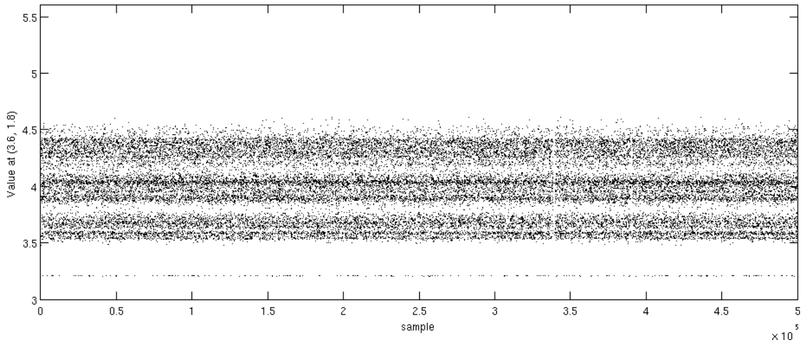
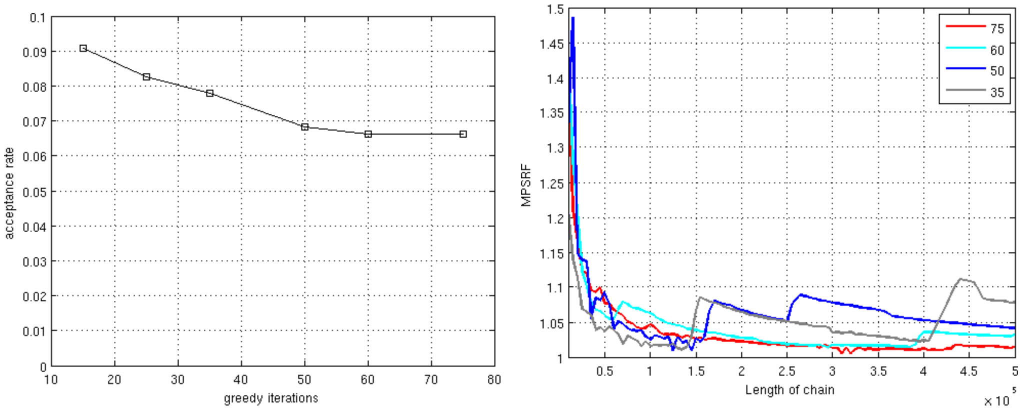


Figure 6.8: Traceplot based on the value of the sample at $(3.6, 1.8)$ for the chain from sampling the posterior pdf with the Gaussian mixture proposal after 60 greedy iterations.

From this chain, we select 10 overdispersed points. We then run 10 chains beginning with these points of 50,000 samples each. We compute the multivariate potential scale reduction factor (MPSRF) [14], which is a measure that compares the covariance within these chains with the overall covariance. Convergence is indicated when the MPSRF approaches 1. The graph in Figure 6.9 plots the MPSRF, and indicates improved convergence after additional greedy iterations. The chain from the Gaussian mixture developed through 75 greedy iterations may be converged well before 500,000 samples, while the chain from the Gaussian mixture developed through 35 greedy iterations, converges much slower.

In Figure 6.10 we plot the error in the estimated mean and variance as a function of greedy iterations, treating the computed mean and variance after 75 greedy iterations as the base case. Both appear to be steadily converging. We then provide a visual guide to the changes in computed mean and variance



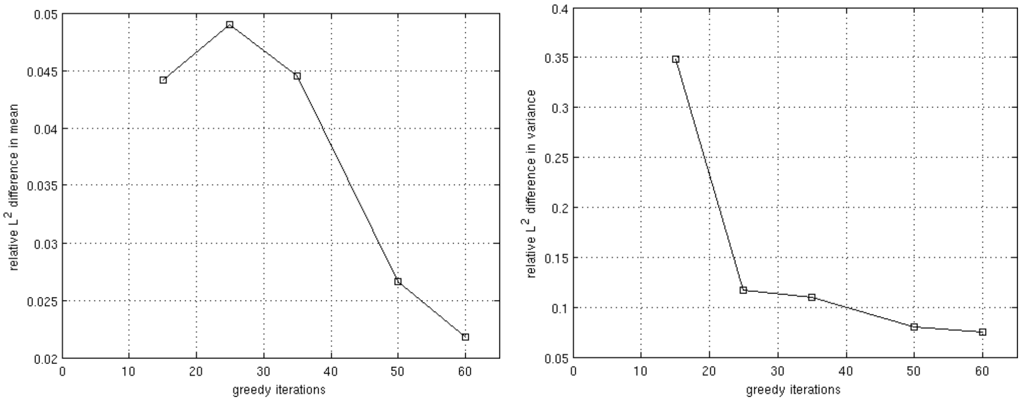
(a) (b)

Figure 6.9: (a) Dependence of the acceptance rate on number of greedy iterations for Metropolis-Hastings sampling of the posterior pdf with the Gaussian mixture proposal. (b) Plots of the MPSRF versus length of chain for Gaussian mixtures built with different numbers of greedy iterations. A value close to 1 suggests convergence.

according to the number of greedy iterations used to build the Gaussian mixture proposal. Most of the changes in the calculated mean (in Figure 6.11) can be seen at the right of the domain and the top left corner. Figure 6.1.2 shows that the variance increases throughout the domain, but especially around the edges.

6.1.3 Metropolis-Hastings sampling of the GPRS with the Gaussian mixture proposal

Next, we consider the results from sampling the GPRS as a surrogate for the posterior pdf, and once again using the Gaussian mixture as a proposal. We first present the traceplot in Figure 6.13 matching the traceplot from sampling the posterior pdf in Figure 6.8. It shows a similar structure, but a much higher



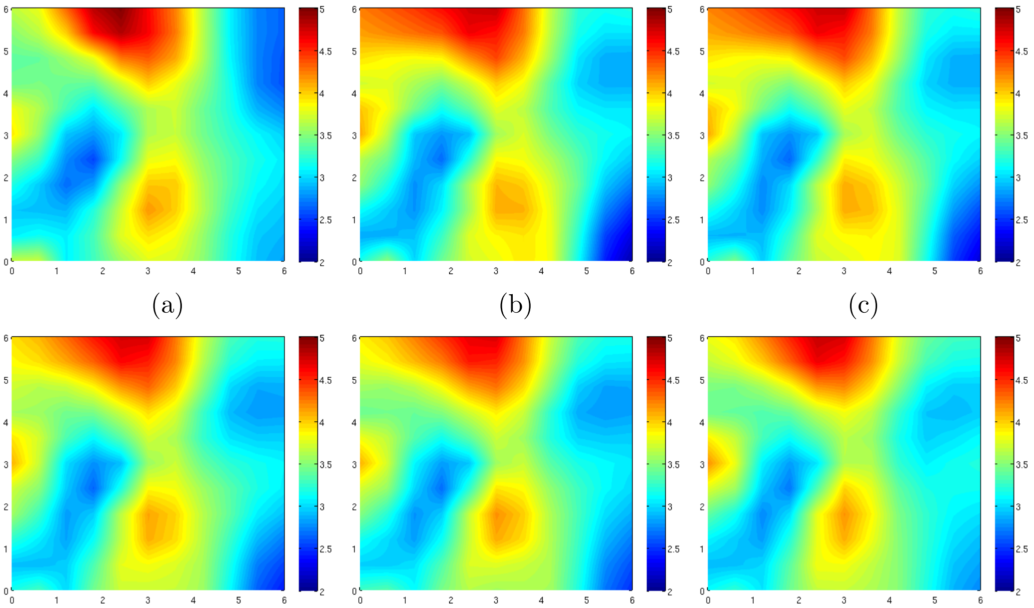
(a) (b)

Figure 6.10: Comparison with mean and variance computed after 75 greedy iterations: (a) Relative L^2 error in the mean versus greedy iterations: $\|m_i - m_{75}\|_{L^2}/\|m_{75}\|_{L^2}$. (b) Relative L^2 error in the variance versus greedy iterations: $\|v_i - v_{75}\|_{L^2}/\|v_{75}\|_{L^2}$

acceptance rate and also a higher variance.

We also plot the MPSRF to investigate convergence rates. Figure 6.1.3 plots the MPSRF and the acceptance rate as a function of greedy iterations. Comparing it with the plots in Figure 6.9, this algorithm has an extremely high acceptance rate and converges very quickly. Since the Gaussian mixture was derived from information extracted from the response surface, it is naturally a good proposal for sampling it.

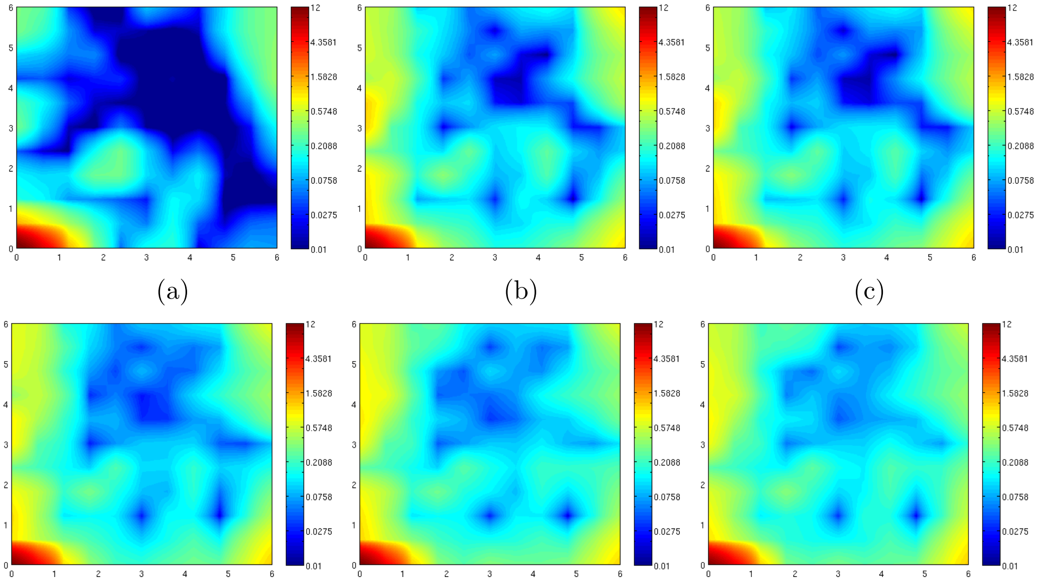
We now compare the mean and variance calculated through sampling the Gaussian process response surface the with the mean and variance from sampling the posterior pdf (where both use the Gaussian mixture as a proposal). Figure 6.15 shows that the mean from the GPRS is a good approximation that improves with the number of greedy iterations. This algorithm



(a) (b) (c)

(d) (e) (f)

Figure 6.11: Mean calculated through Metropolis-Hastings sampling of the posterior pdf with the Gaussian mixture proposal generated by (a) 15 greedy iterations. (b) 25 greedy iterations. (c) 35 greedy iterations. (d) 50 greedy iterations. (e) 60 greedy iterations. (f) 75 greedy iterations.



(a) (b) (c)

(d) (e) (f)

Figure 6.12: Variance calculated through Metropolis-Hastings sampling of the posterior pdf with the Gaussian mixture proposal generated by (a) 15 greedy iterations (b) 25 greedy iterations (c) 35 greedy iterations (d) 50 greedy iterations (e) 60 greedy iterations(f) 75 greedy iterations

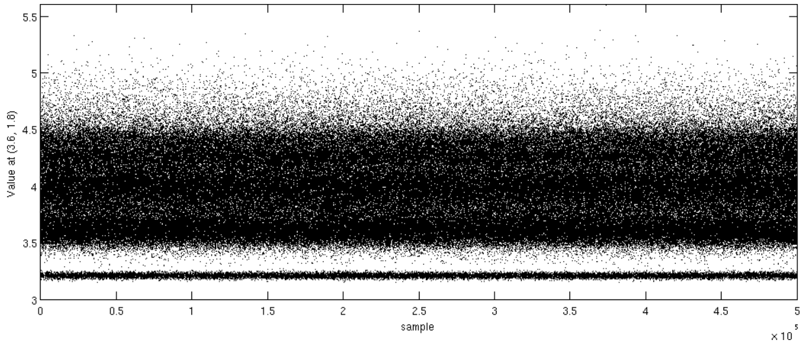
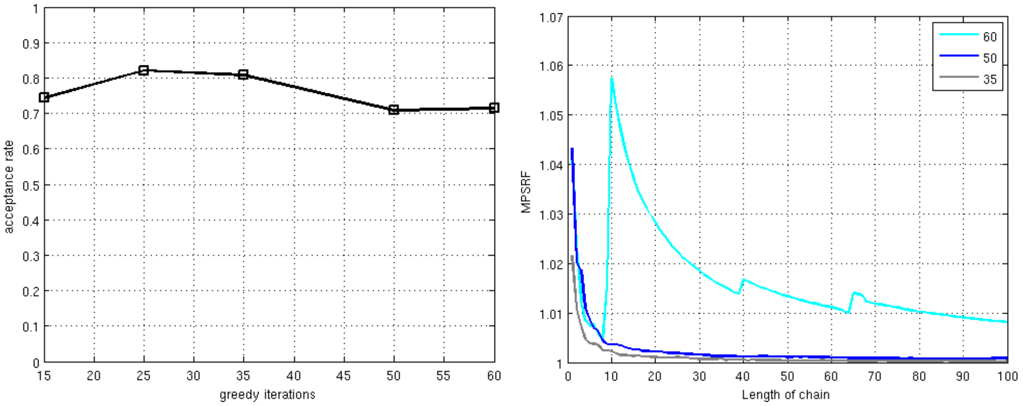


Figure 6.13: Traceplot based on the value of the sample at $(3.6, 1.8)$ for sampling the GPRS with the Gaussian mixture proposal after 60 greedy iterations.



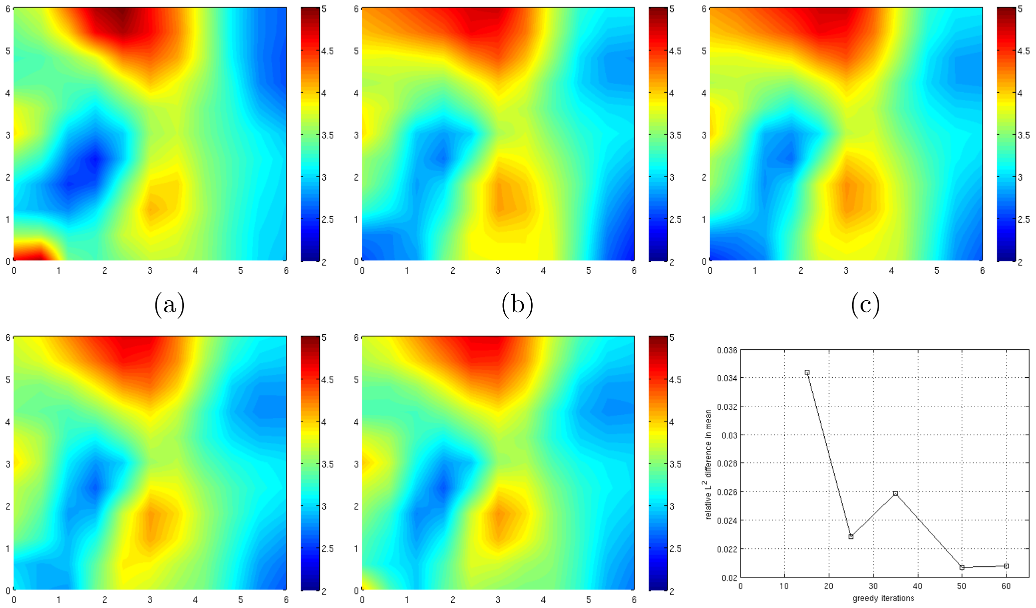
(a) (b)

Figure 6.14: (a) Dependence of the acceptance rate on number of greedy iterations for Metropolis-Hastings sampling of the posterior pdf with the Gaussian mixture proposal. (b) Plots of the MPSRF versus length of chain for Gaussian mixtures built with different numbers of greedy iterations. A value close to 1 suggests convergence.

provides a good approximation of the mean obtained by sampling the posterior pdf at a much lower cost (since no additional PDE solves are required after building the response surface). However, in Figure 6.16 we see that the variance is a poor approximation and getting worse. It is far too large, as was suggested by the traceplot in Figure 6.13. The response surface does not provide a reliable surrogate for the posterior pdf for calculating the variance.

6.1.4 Two-stage delayed acceptance algorithm

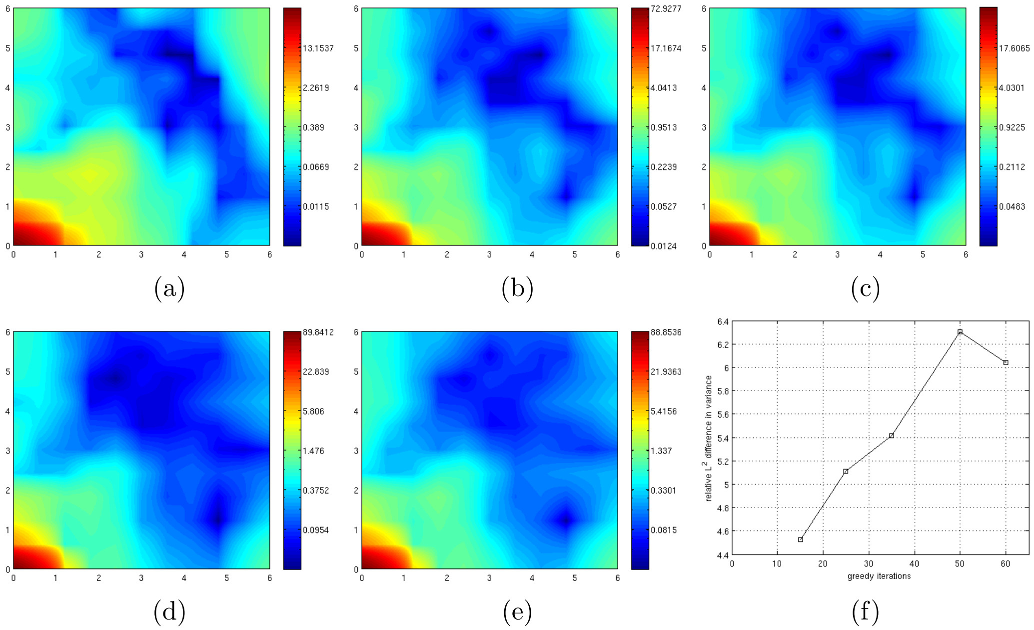
We consider results from the two-stage delayed acceptance algorithm. The traceplot in Figure 6.17 closely resembles that in Figure 6.8. The acceptance rate (see Figure 6.18(a)) is slightly lower. Unfortunately, as can be seen in Figure 6.18(a), the first-stage rejection rate (the rate rejected based on



(a) (b) (c)

(d) (e) (f)

Figure 6.15: Mean calculated through Metropolis-Hastings sampling of the Gaussian process response surface with the Gaussian mixture proposal generated by (a) 15 greedy iterations. (b) 25 greedy iterations. (c) 35 greedy iterations. (d) 50 greedy iterations. (e) 60 greedy iterations. (f) Relative L^2 error in the mean (sampling the response surface versus sampling the posterior pdf) as a function of greedy iterations: $\|m_{\text{gprs}} - m_{\text{posterior}}\|_{L^2} / \|m_{\text{posterior}}\|_{L^2}$.



(a) (b) (c)

(d) (e) (f)

Figure 6.16: Variance calculated through Metropolis-Hastings sampling of the Gaussian process response surface with the Gaussian mixture proposal generated by (a) 15 greedy iterations (b) 25 greedy iterations (c) 35 greedy iterations (d) 50 greedy iterations (e) 60 greedy iterations (f) Relative L^2 error in the mean (sampling the response surface versus sampling the posterior pdf) as a function of greedy iterations: $\|v_{\text{gprs}} - v_{\text{posterior}}\|_{L^2} / \|v_{\text{posterior}}\|_{L^2}$.

the response surface without evaluating the posterior pdf) is low. Given that the acceptance rate when sampling the response surface was very high, this is to be expected, but it makes this algorithm nearly as costly as sampling the posterior pdf directly. The rate at which negative values of the GPRS are encountered (and thus the posterior pdf must be evaluated since we do not trust the response surface) is very high at 60 greedy iterations. This may indicate that more iterations are needed to improve the state of the response surface. We also see in Figure 6.18(b) that it converges more slowly than sampling the posterior pdf.

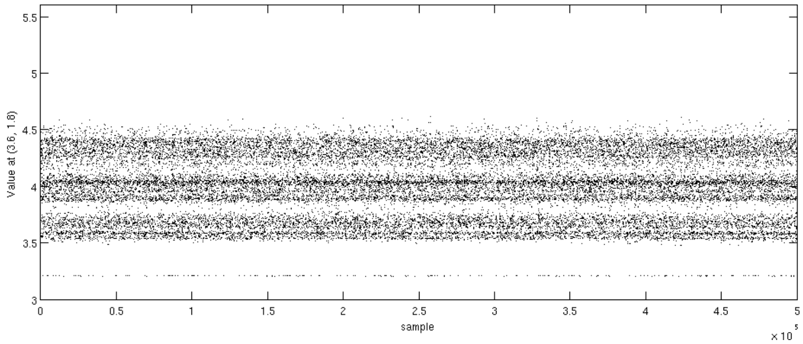
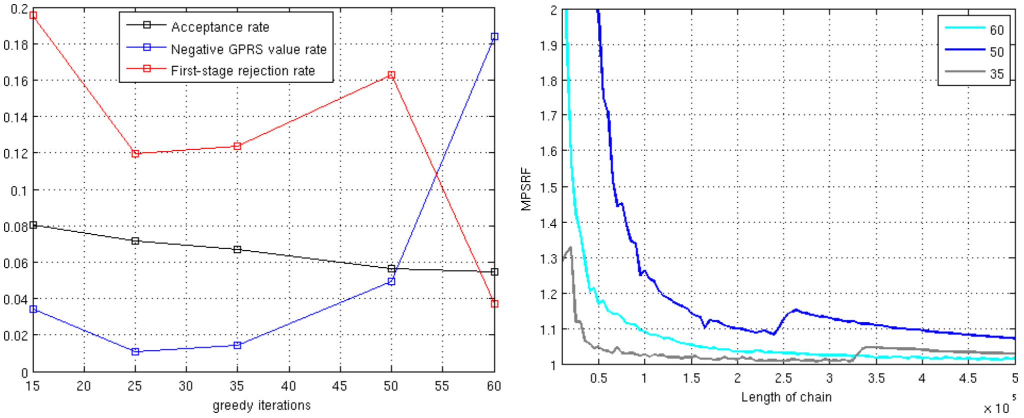


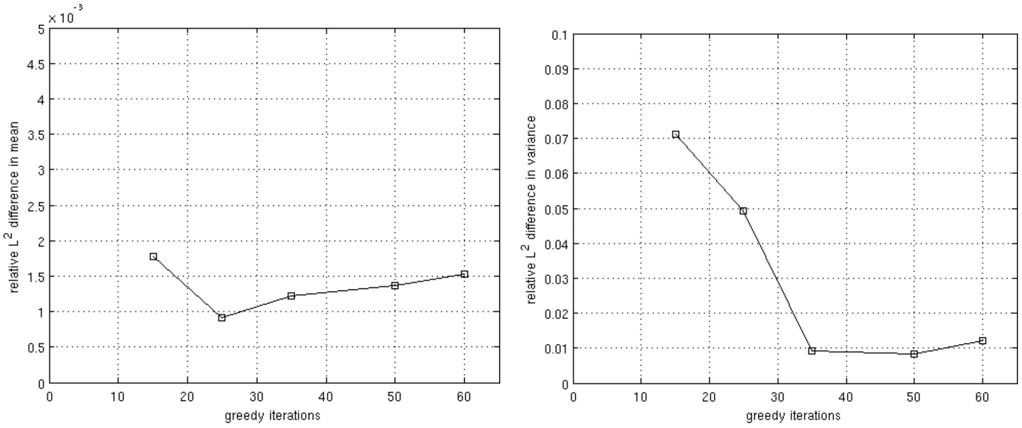
Figure 6.17: Traceplot based on the value of the sample at $(3.6, 1.8)$ for the two-stage delayed acceptance algorithm after 60 greedy iterations.

However, the mean and variance calculated with the two stage delayed acceptance algorithm are very close to the results from sampling the posterior pdf with the Gaussian mixture. The differences are not visible, but we provide graphs of the relative L^2 differences in Figure 6.19. On the other hand, the rejection rate at the first stage is low, so the algorithms are nearly the same regardless.



(a) (b)

Figure 6.18: (a) Dependence of the acceptance rate on number of greedy iterations for Metropolis-Hastings sampling of the posterior pdf with the Gaussian mixture proposal. (b) Plots of the MPSRF versus length of chain for different numbers of greedy iterations. A value close to 1 suggests convergence.



(a) (b)

Figure 6.19: (a) Relative L^2 error in the mean (the two-stage algorithm versus sampling the posterior pdf) as a function of greedy iterations: $\|m_{\text{ts}} - m_{\text{posterior}}\|_{L^2}/\|v_{\text{posterior}}\|_{L^2}$ (b) Relative L^2 error in the variance (the two-stage algorithm versus sampling the posterior pdf) as a function of greedy iterations: $\|v_{\text{ts}} - v_{\text{posterior}}\|_{L^2}/\|v_{\text{posterior}}\|_{L^2}$.

6.1.5 Cost comparison of the three MCMC algorithms

We now compare the three algorithms first proposed in §2.3. We begin with the integrated autocorrelation time in Table 6.1, which is the approximate number of samples in a given algorithm equivalent to obtaining an independent sample of the pdf being sampled. The best is the algorithm sampling the GPRS with the Gaussian mixture. However, sampling the posterior pdf with the Gaussian mixture has an integrated autocorrelation time of around 35, which is very good. The two-stage algorithm is slower.

Algorithm	$(0,0)$	$(3.6,1.8)$	$(0,3)$
1 (75 greedy iterations)	34	31	32
1 (60 greedy iterations)	35	32	32
2 (60 greedy iterations)	2.6	2.1	2.0
3 (60 greedy iterations)	66	39	41

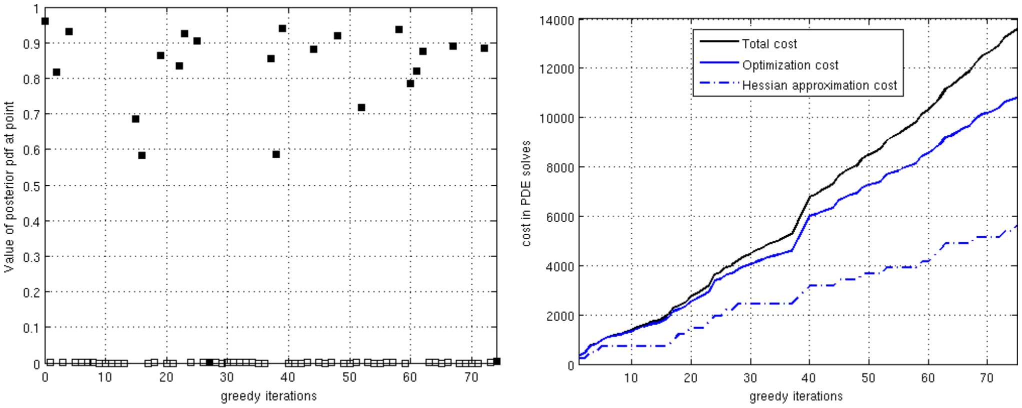
Table 6.1: Integrated autocorrelation time computed based on the value of the sample at the three points listed in the table. Algorithm 1: sampling the posterior pdf with the Gaussian mixture. Algorithm 2: sampling the Gaussian process response surface with the Gaussian mixture. Algorithm 3: the two-stage delayed acceptance algorithm.

Based on the results in Table 6.1, we estimate the costs in PDE solves for the three algorithms, and present them in Table 6.2. They share the upfront cost of building the Hessian-based response surfaces. The cheapest is sampling the Gaussian process response surface as a surrogate, since it requires no additional PDE solves. Recall, however, that it is the least accurate method, and could not provide a good estimate of the variance. The most expensive is the two-stage algorithm. The first-stage rejection rate would need to be considerably higher to offset the need for additional samples.

Algorithm	Upfront cost	Cost per independent sample
1 (75 greedy iterations)	13,594	34
1 (60 greedy iterations)	10,306	35
2 (60 greedy iterations)	10,306	0
3 (60 greedy iterations)	10,306	63

Table 6.2: Cost in PDE solves of the algorithms, based on the maximum of the integrated autocorrelation times in Table 6.1. Algorithm 1: sampling the posterior pdf with the Gaussian mixture. Algorithm 2: sampling the Gaussian process response surface with the Gaussian mixture. Algorithm 3: the two-stage delayed acceptance algorithm. The estimated cost was reduced by the rate of first-stage rejections (which do not require a PDE solve) given in Figure 6.18(a).

In Figure 6.20(a) we present the values of the training points chosen through the greedy algorithm. It is consistently finding training points to improve the response surface, even after 75 greedy iterations. Note the Gaussian mixture after 75 greedy iterations is a mixture of 23 Gaussians. Then in Figure 6.20(b), we present the total cost in PDE solves of building the Hessian-based response surfaces as a function of greedy iterations. Most of the cost comes from the optimization process, not the low-rank Hessian approximations. Every iteration has a full optimization search, but only a third of them build a low-rank Hessian approximation. While this is an expensive algorithm, it produces a good proposal for the posterior pdf (in the form of the Gaussian mixture) and a cheap surrogate for calculating the mean (through sampling the Gaussian process response surface with the Gaussian mixture).



(a) (b)

Figure 6.20: (a) Value of (non-normalized) posterior pdf at training points: Black squares mark shape points (where we include Hessian information) and white squares mark value points (b) Upfront cost of building the Hessian-based response surface in terms of PDE solves.

Chapter 7

Conclusions

Subsurface flow phenomena characterize many important societal issues in energy and the environment. A key feature of these problems is that subsurface properties are uncertain, due to the sparsity of direct observations of the subsurface. The Bayesian formulation of this inverse problem provides a systematic framework for inferring uncertainty in the properties given uncertainties in the data, the forward model, and prior knowledge of the properties. We have addressed the problem: given noisy measurements of the head, the pdf describing the noise, prior information in the form of a pdf of the hydraulic conductivity, and a groundwater flow model relating the head to the hydraulic conductivity, find the posterior probability density function (pdf) of the parameters describing the hydraulic conductivity field. Conventional sampling of this pdf to compute statistical moments is intractable for problems governed by large-scale forward models and high-dimensional parameter spaces.

We constructed a Gaussian process surrogate of the posterior pdf based on Bayesian interpolation between a set of "training" points. We employed a greedy algorithm to find the training points by solving a sequence of optimization problems where each new training point is placed at the maximizer of the error in the approximation. We used scalable Newton optimization methods to solve this “optimal” training point problem. We tailored the Gaussian process surrogate to the curvature of the underlying posterior pdf according to the Hessian of the log posterior at a subset of training points, made computationally tractable by a low-rank approximation of the data misfit Hessian. A Gaussian mixture approximation of the posterior was extracted from the Gaussian process surrogate, and used as a proposal in a Markov chain Monte Carlo method for sampling both the surrogate as well as the true posterior. The Gaussian process surrogate was also used as a first stage approximation in a two-stage delayed acceptance MCMC method.

We provided derivations of expressions for the state, adjoint, gradient, incremental state, incremental adjoint, and Hessian for the groundwater flow problem through calculus of variations with associated discretizations for Galerkin finite elements. We implemented the results in parallel using appropriate scientific libraries.

We verified the suitability and scalability of the low-rank approximation of the prior-preconditioned Hessian of the data misfit through both analysis of infinite dimensional model problems and large-scale numerical results for a convection-diffusion inverse problem and a groundwater flow inverse problem. Future research could include proofs that the Hessian of the data misfit is compact in more general cases than the 1D model problem examples. Analysis of eigenfunctions and eigenvalues may be useful not only for justifying low rank approximations of the Hessian but also point the way to preconditioners for the Hessian.

We implemented the adaptive greedy algorithm to build the Gaussian process response surface described in Chapter 2, and developed a method for generating initial guesses appropriate for high-dimensional spaces. The optimization method used to solve the greedy problem finds the optimum in a mesh-independent number of iterations, and its cost per iteration requires a number of PDE solves that is independent of problem dimension (when the low rank approximation is employed). The cost of evaluating either the Gaussian mixture or the Gaussian process response surface are negligible compared to the cost of solving a PDE for the forward problem for a large-scale problem.

We also developed a Gaussian mixture response surface, which our numerical evidence suggests is a good proposal density for the posterior pdf in high-dimensional parameter spaces. In our comparison in Chapter 6, a state-of-the-art reference method (DRAM) failed to produce a usable chain after 1,000,000 samples (at a cost of 1 PDE solve each) for a problem with 121 parameters. The Gaussian mixture proposal had an upfront cost of between 10,000 and 14,000 PDE solves (depending on the choice of cutoff for the number of greedy iterations) to begin producing a chain with a cost of only approximately 35 PDE solves per additional independent sample of the posterior pdf. Further research in choosing the best weights for the mixture may improve performance even more.

We proposed and implemented three MCMC algorithms using the Hessian-based response surfaces and assessed their convergence, accuracy, and cost for a groundwater flow inverse problem. We also provided results and analysis of properties of the posterior pdf for both the convection-diffusion inverse problem and the groundwater flow inverse problem. In the future, the algorithms described here could be applied to larger, more complex problems such as coupled flow and transport models, or 3D groundwater flow problems.

The Gaussian process response surface proved a useful surrogate for the posterior pdf when using Metropolis-Hastings with the Gaussian mixture proposal to calculate the mean of the posterior pdf at a greatly reduced cost. However, it is a poor surrogate for other properties of the posterior pdf, such as the variance. Future research could include investigating alternative interpolation prior covariance functions that enforce the curvature of the Hessian but include additional conditions such as positivity.

While the idea behind the two-stage delayed acceptance algorithm was promising, in practice it increased costs since it did not reject enough points in the first stage. However, with further research to improve the quality of the Gaussian process response surface as a first-stage filter, it could still prove useful. Since the rejection rate from using the Gaussian mixture proposal to sample the posterior pdf directly is still above 90%, there is a lot of room for improvement.

Bibliography

[1] Volkan Akçelik, George Biros, Andrei Draganescu, Omar Ghattas, Judith Hill, and Bart van Bloeman Waanders. Dynamic data-driven inversion for terascale simulations: Real-time identification of airborne contaminants. In *Proceedings of SC2005*, Seattle, 2005.

[2] Volkan Akçelik, George Biros, Omar Ghattas, Kevin R. Long, and Bart van Bloemen Waanders. A variational finite element method for source inversion for convective-diffusive transport. *Finite Elements in Analysis and Design*, 39(8):683–705, 2003.

[3] G. Backus and F. Gilbert. The resolving power of gross earth data. *Geophysical Journal of the Royal Astronomical Society*, 16:169–205, 1968.

[4] S. Balakrishnan, A. Roy, M. G. Ierapetritou, G. P. Flach, and P. G. Georgopoulos. Uncertainty reduction and characterization for complex environmental fate and transport models: An empirical bayesian framework incorporating the stochastic response surface method. *Water Resources Research*, 39(12):8–1–8–13, December 2003.

[5] S. Balay, K. Buschelman, V. Eijkhout, W. D. Gropp, D. Kaushik, M. G. Knepley, L. C. McInnes, B. F. Smith, and H. Zhang. PETSc Users

Manual. Technical Report ANL-95/11 - Revision 3.0.0, Argonne National Laboratory, 2008.

[6] S. Balay, K. Buschelman, W. D. Gropp, D. Kaushik, M. G. Knepley, L. C. McInnes, B. F. Smith, and H. Zhang. PETSc Web page, 2009. http://www.mcs.anl.gov/petsc.

[7] Satish Balay, William D. Gropp, Lois Curfman McInnes, and Barry F. Smith. Efficient management of parallelism in object oriented numerical software libraries. In E. Arge, A. M. Bruaset, and H. P. Langtangen, editors, *Modern Software Tools in Scientific Computing*, pages 163–202. Birkhäuser Press, 1997.

[8] Wolfgang Bangerth, Ralf Hartmann, and Guido Kanschat. `deal.II` *Differential Equations Analysis Library, Technical Reference.* Interdisziplinäres Zentrum für Wissenschaftliches Rechnen (IWR). `http://www.dealii.org`.

[9] Wolfgang Bangerth, Ralf Hartmann, and Guido Kanschat. deal.II – a general-purpose object-oriented finite element library. *ACM Transactions on Mathematical Software*, 33(4):24, 2007.

[10] Jacob Bear. *Dynamics of fluids in porous media.* Dover Publications, Inc., 1972.

[11] J. G. Berryman. Analysis of approximate inverses in tomography I. Resolution analysis of common inverses. *Optimization and Engineering*, 1(1):87–115, 2001.

[12] J. G. Berryman. Analysis of approximate inverses in tomography II. Iterative inverses. *Optimization and Engineering*, 1(1):437–473, 2001.

[13] L. Boschi. Measures of resolution in global body wave tomography. *Geophysical Research Letters*, 30(19), 2003.

[14] Stephen Brooks and Andrew Gelman. General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*, 7:434–456, 1998.

[15] Tan Bui-Thanh, Omar Ghattas, and David Higdon. Adaptive Hessian-based nonstationary Gaussian process response surface method for probability density approximation with application to Bayesian solution of large-scale inverse problems. *SIAM Journal on Scientific Computing*, 34(6):A2837–A2871, 2012.

[16] Tan Bui-Thanh, Omar Ghattas, James Martin, and Georg Stadler. A computational framework for infinite-dimensional Bayesian inverse problems. Part I: The linearized case, with applications to global seismic inversion, 2012. Submitted.

[17] Tan Bui-Thanh, Karen Willcox, and Omar Ghattas. Model reduction for large-scale systems with high-dimensional parametric input space. *SIAM Journal on Scientific Computing*, 30:3270–3288, 2008.

[18] Andres Christen and Bruno Sanso. Advances in the design of Gaussian processes as surrogate models for computer experiments. Technical

Report 5, University of California, Santa Cruz, 2008.

[19] J. Andrés Christen and Colin Fox. Markov chain Monte Carlo using an approximation. *Journal of Computational & Graphical Statistics*, 14(4):795–810, December 2005.

[20] David A. Cohn. Neural networks exploration using optimal experiment design. *Neural Networks*, 9(6):1071–1083, 1996.

[21] T. Cui, C. Fox, and M. J. O'Sullivan. Bayesian calibration of a large-scale geothermal reservoir model by a new adaptive delayed acceptance Metropolis Hastings algorithm. *Water Resources*, 47(10), 2011.

[22] P. Dostert, Y. Efendiev, and B. Mohanty. Efficient uncertainty quantification techniques in inverse problems for Richards' equation using coarse-scale simulation models. *Advances in Water Resources*, 32:329–339, 2009.

[23] Paul Dostert, Yalchin Efendiev, Thomas Y. Hou, and Wuan Luo. Coarse-gradient Langevin algorithms for dynamic data integration and uncertainty quantification. *Journal of Computational Physics*, 217:123–142, 2006.

[24] Y. Efendiev, A. Datta-Gupta, V. Ginting, X. Ma, and B. Mallick. An efficient two-stage Markov Chain Monte Carlo method for dynamic data integration. *Water Resources Research*, 41, 2004.

[25] Y. Efendiev, T. Hou, and W. Luo. Preconditioning Markov Chain Monte Carlo simulations using coarse-scale models. *SIAM Journal on Scientific Computing*, 28(2):776–803, 2006.

[26] H. Pearl Flath, Lucas C. Wilcox, Volkan Akçelik, Judy Hill, Bart van Bloemen Waanders, and Omar Ghattas. Fast algorithms for Bayesian uncertainty quantification in large-scale linear inverse problems based on low-rank partial Hessian approximations. *SIAM Journal on Scientific Computing*, 33(1):407–432, 2011.

[27] S. Fomel, J. Berryman, R. Clapp, and M.F. Prucha. Iterative resolution estimation in least squares Kirchhoff migration. *Geophysical Prospecting*, 50:577–588, 2002.

[28] W. P. Gouveia and J. A. Scales. Resolution of seismic waveform inversion: Bayes versus Occam. *Inverse Problems*, 13:323–349, 1997.

[29] W. P. Gouveia and J. A. Scales. Bayesian seismic waveform inversion: Parameter estimation and uncertainty analysis. *Journal of Geophysical Research*, 103(B2):2759–2779, 1998.

[30] Heikki Haario, Marko Laine, Antonietta Miravete, and Eero Saksman. DRAM: Efficient adaptive MCMC. *Statistics and Computing*, 16:339–354, 2006.

[31] M. Hanke. A regularizing levenberg-marquardt scheme, with applications to inverse groundwater filtration problems. *Inverse Prob*, 13:79–95, 1997.

[32] V. Hernandez, J. E. Roman, A. Tomas, and V. Vidal. Lanczos methods in SLEPc. Technical Report STR-5, Universidad Politécnica de Valencia, 2006. Available at http://www.grycap.upv.es/slepc.

[33] V. Hernandez, J. E. Roman, and V. Vidal. SLEPc: A scalable and flexible toolkit for the solution of eigenvalue problems. *ACM Transactions on Mathematical Software*, 31(3):351–362, September 2005.

[34] David Higdon, Marc Kennedy, James C. Cavendish, John A. Cafeo, and Robert D. Ryne. Combining field data and computer simulations for calibration and prediction. *SIAM Journal on Scientific Computing*, 26(2):448–466, 2004.

[35] David Higdon, Herbis Lee, and Chris Holloman. Markov chain Monte Carlo-based approaches for inference in computationally intensive inverse problems. In J. M. Bernardo, M. J. Bayarri, J. O. Berger, A. P. Dawid, D. Heckerman, A. F. M. Smith, and M. West, editors, *Bayesian Statistics 7*, 2003.

[36] Marco Antonio Iglesias-Hernandez. *An Iterative Representer-Based Scheme for Data Inversion in Reservoir Modeling.* PhD thesis, The University of Texas at Austin, 2008.

[37] Bangti Jin. Fast Bayesian approach for parameter estimation. *International Journal for Numerical Methods in Engineering*, 76(2):230 – 252, 2008.

[38] Jari Kaipio and Erkki Somersalo. *Statistical and Computational Inverse Problems*, volume 160 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 2005.

[39] Marc C. Kennedy and Anthony O'Hagan. Bayesian calibration of computer models. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 63(3):425–464, 2001.

[40] C. Lieberman, K. Fidkowski, K. Willcox, and O. Ghattas. Real-time data assimilation using Hessian-based model reduction: Application to linear contaminant transport. In preparation, 2009.

[41] Chad Lieberman, Karen Willcox, and Omar Ghattas. Parameter and state model reduction for large-scale statistical inverse problems. *SIAM Journal on Scientific Computing*, 32(5):2523–2542, 2010.

[42] O. Marques, T. Drummond, and D. Vasco. A computational strategy for the solution of large linear inverse problems in geophysics. *Parallel and Distributed Processing Symposium, International*, page 16b, 2003.

[43] James Martin, Lucas C. Wilcox, Carsten Burstedde, and Omar Ghattas. A stochastic Newton MCMC method for large-scale statistical inverse problems with application to seismic inversion. *SIAM Journal on Scientific Computing*, 34(3):A1460–A1487, 2012.

[44] Y. Marzouk and D. Xiu. A stochastic collocation approach to Bayesian inference in inverse problems. *Communications in Computational Physics*,

6(4):826–847, 2009.

[45] Youssef M. Marzouk and Habib N. Najm. Dimensionality reduction and polynomial chaos acceleration of Bayesian inference in inverse problems. *Journal of Computational Physics*, 228:1862–1902, 2009.

[46] S. Minkoff. A computationally feasible approximate resolution matrix for seismic inverse problems. *Geophysical Journal International*, 126:345–359, 1996.

[47] G. Nolet, R. Montelli, and J. Virieux. Explicit, approximate expressions for the resolution and a-posteriori covariance of massive tomographic systems. *Geophysical Journal International*, 138:36–44, 1999.

[48] Dean S. Oliver, Albert C. Reynolds, and Ning Liu. *Inverse theory for petroleum reservoir characterization and history matching.* Cambidge University Press, 2008.

[49] N. Petra, J. Martin, G. Stadler, and O. Ghattas. A computational framework for infinite-dimensional Bayesian inverse problems: Part II. Stochastic Newton MCMC with application to ice sheet inverse problems. *In preparation*, 2013.

[50] Carl E. Rasmussen and Christopher K. Williams. *Gaussian Processes for Machine Learning.* MIT Press, 2006.

[51] Andrew M. Stuart. Inverse problems: A Bayesian perspective. *Acta Numerica*, 19:451–559, 2010.

[52] Albert Tarantola. *Inverse Problem Theory and Methods for Model Parameter Estimation*. SIAM, Philadelphia, PA, 2005.

[53] D. W. Vasco, Jr. J. E. Peterson, and E. L. Majer. Resolving seismic anisotropy: Sparse matrix methods for geophysical inverse problems. *Geophysics*, 63(3):970–983, 1998.

[54] D. W. Vasco, R. J. Pulliam, and L. R. Johnson. Formal inversion of ISC arrival times for mantle P-velocity structure. *Geophysical Journal International*, 113:586–606, 1993.

[55] Curt R. Vogel. *Computational Methods for Inverse Problems*. Frontiers in Applied Mathematics. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002.

[56] Jingbo Wang and Nicholas Zabaras. Hierarchical Bayesian models for inverse problems in heat conduction. *Inverse Problems*, 21(1):183–206, 2005.

[57] Z. S. Yao, R. G. Roberts, and A. Tryggvason. Calculating resolution and covariance matrices for seismic tomography with the LSQR method. *Geophysical Journal International*, 138:886–894, 1999.

[58] H. Zhang and C. H. Thurber. Estimating the model resolution matrix for large seismic tomography problems based on Lanczos bidiagonalization with partial reorthogonalization. *Geophysical Journal International*, 170:337–345, 2007.

[59] J. Zhang and G. A. McMechan. Estimation of resolution and covariance for large matrix inversions. *Geophysical Journal International*, 121:409–426, 1995.

Vita

Pearl Flath received a Ph.D. in Computational Science, Engineering, and Mathematics in August 2013 from The University of Texas at Austin, where she also earned an M.S. in Computational and Applied Mathematics. She has a B.A. in Applied Mathematics and a concentration in Physics from Columbia University. Her research focuses on uncertainty quantification in large scale statistical inverse problems. She was awarded the Department of Homeland Security Graduate Fellowship and the CAM Fellowship, and was an I.I. Rabi Scholar at Columbia University. She is active in SIAM, serving as founding president of the SIAM student chapter at both The University of Texas at Austin and Columbia University.

Permanent address: 3456 North Hills Dr, Apt 319
Austin, TX 78731

This dissertation was typeset with LaTeX[†] by the author.

[†] LaTeX is a document preparation system developed by Leslie Lamport as a special version of Donald Knuth's TeX Program.